

PATENT
4,100-0446

10/509278

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

DT04 Rec'd PCT/PTO 27 SEP 2004

Applicants: Noriyuki YAMAMOTO et al.

International Application No.: PCT/JP03/03795

International Filing Date: March 27, 2003

For: INFORMATION SEARCH SYSTEM, INFORMATION
PROCESSING APPARATUS AND METHOD, AND
INFORMATION SEARCH APPARATUS AND
METHOD

745 Fifth Avenue
New York, NY 10151

EXPRESS MAIL

Mailing Label Number: EV385414560US

Date of Deposit: September 27, 2004

I hereby certify that this paper or fee is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" Service under 37 CFR 1.10 on the date indicated above and is addressed to Mail Stop PCT, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.

Barnet Shindlmayr
(Typed or printed name of person mailing paper or fee)

Barry Bell
(Signature of person mailing paper or fee)

CLAIM OF PRIORITY UNDER 37 C.F.R. § 1.78(a)(2)

Mail Stop PCT
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Sir:

Pursuant to 35 U.S.C. 119, this application is entitled to a claim of priority to Japan Application No. 2002-095414 filed 29 March 2002.

Respectfully submitted,

FROMMER LAWRENCE & HAUG LLP
Attorneys for Applicants

By:

William Frommer
William S. Frommer
Reg. No. 25,506
Tel. (212) 588-0800

10/509278

PCT/JP03/03795
Rec'd PCT/PTO 27 SEP 2004

日本国特許庁
JAPAN PATENT OFFICE

27.03.03

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

出願年月日

Date of Application:

2002年 3月29日

REC'D 17 APR 2003

出願番号

Application Number:

特願2002-095414

WIPO PCT

[ST.10/C]:

[J P 2002-095414]

出願人

Applicant(s):

ソニー株式会社

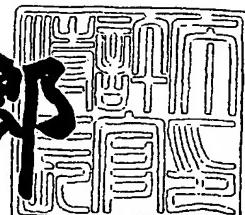
PRIORITY
DOCUMENT

SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH RULE 17.1 (a) OR (b)

2002年11月29日

特許庁長官
Commissioner,
Japan Patent Office

太田 信一郎



出証番号 出証特2002-3094555

Best Available Copy

【書類名】 特許願
 【整理番号】 0290199314
 【提出日】 平成14年 3月29日
 【あて先】 特許庁長官殿
 【国際特許分類】 G06F 19/00
 【発明者】
 【住所又は居所】 東京都品川区北品川6丁目7番35号 ソニー株式会社
 内
 【氏名】 山本 則行
 【発明者】
 【住所又は居所】 東京都品川区北品川6丁目7番35号 ソニー株式会社
 内
 【氏名】 斎藤 真里
 【特許出願人】
 【識別番号】 000002185
 【氏名又は名称】 ソニー株式会社
 【代理人】
 【識別番号】 100082131
 【弁理士】
 【氏名又は名称】 稲本 義雄
 【電話番号】 03-3369-6479
 【手数料の表示】
 【予納台帳番号】 032089
 【納付金額】 21,000円
 【提出物件の目録】
 【物件名】 明細書 1
 【物件名】 図面 1
 【物件名】 要約書 1
 【包括委任状番号】 9708842

特2002-095414

【ブルーフの要否】 要

【書類名】 明細書

【発明の名称】 情報検索システム、情報処理装置および方法、情報検索装置および方法、記録媒体、並びにプログラム

【特許請求の範囲】

【請求項1】 情報処理装置、およびネットワークを介して前記情報処理装置からアクセスされる情報検索装置からなる情報検索システムにおいて、

前記情報処理装置は、

所定の情報を分析し、番組に関する番組情報を取得するための興味語を抽出する抽出手段と、

前記抽出手段により抽出された前記興味語を前記情報検索装置に送信し、前記興味語に対応する前記番組情報の検索を要求する検索要求手段と、

前記検索要求手段に基づいて、前記情報検索装置から送信されてくる前記番組情報を受信する受信手段と

を備え、

前記情報検索装置は、

前記番組情報を蓄積する蓄積手段と、

前記情報処理装置から送信されてきた前記検索要求に基づいて、前記検索要求に含まれる前記興味語に関連する前記番組情報を前記蓄積手段から検索する検索手段と、

前記検索手段により検索された前記番組情報を前記情報処理装置に送信する送信手段と

を備えることを特徴とする情報検索システム。

【請求項2】 前記情報処理装置の前記抽出手段は、前記所定の情報を形態素解析して前記興味語に分解する形態素解析手段を含む

ことを特徴とする請求項1に記載の情報検索システム。

【請求項3】 前記情報処理装置は、前記抽出手段により抽出された前記興味語を用いてデータベースを作成するデータベース構築手段をさらに備えることを特徴とする請求項1に記載の情報検索システム。

【請求項4】 前記情報処理装置は、前記受信手段により受信された前記番

組情報に基づいて、前記番組の録画を制御する録画制御手段をさらに備えることを特徴とする請求項1に記載の情報検索システム。

【請求項5】 前記情報処理装置は、前記受信手段により受信された前記番組情報の表示を制御する表示制御手段をさらに備えることを特徴とする請求項1に記載の情報検索システム。

【請求項6】 前記情報検索装置の前記蓄積手段は、前記番組情報を用いて、前記番組に関連付けてデータベースを作成するデータベース構築手段を含むことを特徴とする請求項1に記載の情報検索システム。

【請求項7】 前記所定の情報は、文書情報、前記番組に関する嗜好情報、または、前記番組の視聴履歴に関する情報のうち、少なくとも1つを含むことを特徴とする請求項1に記載の情報検索システム。

【請求項8】 前記文書情報は、電子メールであることを特徴とする請求項7に記載の情報検索システム。

【請求項9】 前記番組情報は、前記番組を録画するための、録画開始時刻、録画終了時刻、およびチャンネル情報を含むことを特徴とする請求項1に記載の情報検索システム。

【請求項10】 所定の情報を分析し、番組に関する番組情報を取得するための興味語を抽出する抽出手段と、

前記抽出手段により抽出された前記興味語を情報検索装置に送信し、前記興味語に対応する前記番組情報の検索を要求する検索要求手段と、

前記検索要求手段に基づいて、前記情報検索装置から送信されてくる前記番組情報を受信する受信手段と

を備えることを特徴とする情報処理装置。

【請求項11】 前記抽出手段は、前記所定の情報を形態素解析して前記興味語に分解する形態素解析手段を含むことを特徴とする請求項10に記載の情報処理装置。

【請求項12】 前記抽出手段により抽出された前記興味語を用いてデータベースを作成するデータベース構築手段をさらに備えることを特徴とする請求項10に記載の情報処理装置。

【請求項13】 前記受信手段により受信された前記番組情報に基づいて、前記番組の録画を制御する録画制御手段をさらに備えることを特徴とする請求項10に記載の情報処理装置。

【請求項14】 前記受信手段により受信された前記番組情報の表示を制御する表示制御手段をさらに備えることを特徴とする請求項10に記載の情報処理装置。

【請求項15】 前記所定の情報は、文書情報、前記番組に関する嗜好情報、または、前記番組の視聴履歴に関する情報のうち、少なくとも1つを含むことを特徴とする請求項10に記載の情報処理装置。

【請求項16】 前記文書情報は、電子メールであることを特徴とする請求項15に記載の情報処理装置。

【請求項17】 前記番組情報は、前記番組を録画するための、録画開始時刻、録画終了時刻、およびチャンネル情報を含むことを特徴とする請求項10に記載の情報処理装置。

【請求項18】 番組の録画を制御する録画制御手段を備える情報処理装置の情報処理方法において、

前記所定の情報を分析し、番組に関する番組情報を取得するための興味語を抽出する抽出ステップと、

前記抽出ステップの処理により抽出された前記興味語を情報検索装置に送信し、前記興味語に対応する前記番組情報の検索を要求する検索要求ステップと、前記検索要求ステップの処理に基づいて、前記情報検索装置から送信されてくる前記番組情報の受信を制御する受信制御ステップとを含むことを特徴とする情報処理方法。

【請求項19】 番組の録画を制御する録画制御手段を備える情報処理装置を制御するプログラムであって、

前記所定の情報を分析し、番組に関する番組情報を取得するための興味語を抽出する抽出ステップと、

前記抽出ステップの処理により抽出された前記興味語を情報検索装置に送信し、前記興味語に対応する前記番組情報の検索を要求する検索要求ステップと、

前記検索要求ステップの処理に基づいて、前記情報検索装置から送信されてくる前記番組情報の受信を制御する受信制御ステップと

を含むことを特徴とするコンピュータが読み取り可能なプログラムが記録されている記録媒体。

【請求項20】 番組の録画を制御する録画制御手段を備える情報処理装置を制御するコンピュータに、

前記所定の情報を分析し、番組に関する番組情報を取得するための興味語を抽出する抽出ステップと、

前記抽出ステップの処理により抽出された前記興味語を情報検索装置に送信し、前記興味語に対応する前記番組情報の検索を要求する検索要求ステップと、

前記検索要求ステップの処理に基づいて、前記情報検索装置から送信されてくる前記番組情報の受信を制御する受信制御ステップと

を実行させるプログラム。

【請求項21】 番組に関する番組情報を蓄積する蓄積手段と、

情報処理装置から送信されてきた、前記番組情報を取得するための興味語を受信する受信手段と、

前記受信手段により受信された前記興味語に関連する前記番組情報を前記蓄積手段から検索する検索手段と、

前記検索手段により検索された前記番組情報を前記情報処理装置に送信する送信手段と

を備えることを特徴とする情報検索装置。

【請求項22】 前記興味語は、前記情報処理装置において所定の情報が形態素解析されて分解された単語である

ことを特徴とする請求項21に記載の情報検索装置。

【請求項23】 前記番組情報は、前記番組を録画するための、録画開始時刻、録画終了時刻、およびチャンネル情報を含む

ことを特徴とする請求項21に記載の情報検索装置。

【請求項24】 情報を検索する情報検索装置の情報検索方法において、番組に関する番組情報の蓄積を制御する蓄積制御ステップと、

情報処理装置から送信されてきた、前記番組情報を取得するための興味語の受信を制御する受信制御ステップと、

前記受信制御ステップの処理により受信が制御された前記興味語に関連する前記番組情報を検索する検索ステップと、

前記検索ステップの処理により検索された前記番組情報の、前記情報処理装置への送信を制御する送信制御ステップと

を含むことを特徴とする情報検索方法。

【請求項25】 情報を検索する情報検索装置を制御するプログラムであつて、

番組に関する番組情報の蓄積を制御する蓄積制御ステップと、

情報処理装置から送信されてきた、前記番組情報を取得するための興味語の受信を制御する受信制御ステップと、

前記受信制御ステップの処理により受信が制御された前記興味語に関連する前記番組情報を検索する検索ステップと、

前記検索ステップの処理により検索された前記番組情報の、前記情報処理装置への送信を制御する送信制御ステップと

を含むことを特徴とするコンピュータが読み取り可能なプログラムが記録されている記録媒体。

【請求項26】 情報を検索する情報検索装置を制御するコンピュータに、
番組に関する番組情報の蓄積を制御する蓄積制御ステップと、
情報処理装置から送信されてきた、前記番組情報を取得するための興味語の受信を制御する受信制御ステップと、

前記受信制御ステップの処理により受信が制御された前記興味語に関連する前記番組情報を検索する検索ステップと、

前記検索ステップの処理により検索された前記番組情報の、前記情報処理装置への送信を制御する送信制御ステップと
を実行させるプログラム。

【請求項27】 携帯端末装置、前記携帯端末装置にネットワークを介して接続される情報処理装置、および、前記ネットワークを介して前記情報処理装置

からアクセスされる情報検索装置からなる情報検索システムにおいて、

前記携帯端末装置は、

番組を録画予約するための予約情報を生成する生成手段と、

前記生成手段により生成された前記予約情報を前記情報処理装置に送信する第1の送信手段と

を備え、

前記情報処理装置は、

前記携帯端末装置から送信されてきた前記予約情報を分析し、前記番組に関する番組情報を取得するための興味語を抽出する抽出手段と、

前記抽出手段により抽出された前記興味語を前記情報検索装置に送信し、前記興味語に対応する前記番組情報の検索を要求する検索要求手段と、

前記検索要求手段に基づいて、前記情報検索装置から送信されてくる前記番組情報を受信する受信手段と

を備え、

前記情報検索装置は、

前記番組情報を蓄積する蓄積手段と、

前記情報処理装置から送信されてきた前記検索要求に基づいて、前記検索要求に含まれる前記興味語に関連する前記番組情報を前記蓄積手段から検索する検索手段と、

前記検索手段により検索された前記番組情報を前記情報処理装置に送信する第2の送信手段と

を備えることを特徴とする情報検索システム。

【請求項28】 前記情報処理装置の前記抽出手段は、前記所定の情報を形態素解析して前記興味語に分解する形態素解析手段を含む

ことを特徴とする請求項27に記載の情報検索システム。

【請求項29】 前記情報処理装置は、前記受信手段により受信された前記番組情報に基づいて、前記番組の録画を制御する録画制御手段をさらに備えることを特徴とする請求項27に記載の情報検索システム。

【請求項30】 前記情報検索装置の前記蓄積手段は、前記番組情報を用い

て、前記番組に関連付けてデータベースを作成するデータベース構築手段を含むことを特徴とする請求項27に記載の情報検索システム。

【請求項31】 前記予約情報は、番組名、ジャンル名、または、出演者名のうち、少なくとも1つを含むことを特徴とする請求項27に記載の情報検索システム。

【請求項32】 前記予約情報は、電子メールであることを特徴とする請求項27に記載の情報検索システム。

【請求項33】 前記番組情報は、前記番組を録画するための、録画開始時刻、録画終了時刻、およびチャンネル情報を含むことを特徴とする請求項27に記載の情報検索システム。

【請求項34】 携帯端末装置から送信されてきた、番組を録画予約するための予約情報を分析し、前記番組に関する番組情報を取得するための興味語を抽出する抽出手段と、

前記抽出手段により抽出された前記興味語を情報検索装置に送信し、前記興味語に対応する前記番組情報の検索を要求する検索要求手段と、

前記検索要求手段に基づいて、前記情報検索装置から送信されてくる前記番組情報を受信する受信手段と

を備えることを特徴とする情報処理装置。

【請求項35】 前記抽出手段は、前記所定の情報を形態素解析して前記興味語に分解する形態素解析手段を含むことを特徴とする請求項34に記載の情報処理装置。

【請求項36】 前記受信手段により受信された前記番組情報に基づいて、前記番組の録画を制御する録画制御手段をさらに備えることを特徴とする請求項34に記載の情報処理装置。

【請求項37】 前記予約情報は、番組名、ジャンル名、または、出演者名のうち、少なくとも1つを含むことを特徴とする請求項34に記載の情報処理装置。

【請求項38】 前記予約情報は、電子メールであることを特徴とする請求項34に記載の情報処理装置。

【請求項39】 前記番組情報は、前記番組を録画するための、録画開始時刻、録画終了時刻、およびチャンネル情報を含むことを特徴とする請求項34に記載の情報処理装置。

【請求項40】 番組の録画を制御する録画制御手段を備える情報処理装置の情報処理方法において、

携帯端末装置から送信されてきた、番組を録画予約するための予約情報を分析し、前記番組に関する番組情報を取得するための興味語を抽出する抽出ステップと、

前記抽出ステップの処理により抽出された前記興味語を情報検索装置に送信し、前記興味語に対応する前記番組情報の検索を要求する検索要求ステップと、前記検索要求ステップの処理に基づいて、前記情報検索装置から送信されてくる前記番組情報の受信を制御する受信制御ステップと

を含むことを特徴とする情報処理方法。

【請求項41】 番組の録画を制御する録画制御手段を備える情報処理装置を制御するプログラムであって、

携帯端末装置から送信されてきた、番組を録画予約するための予約情報を分析し、前記番組に関する番組情報を取得するための興味語を抽出する抽出ステップと、

前記抽出ステップの処理により抽出された前記興味語を情報検索装置に送信し、前記興味語に対応する前記番組情報の検索を要求する検索要求ステップと、前記検索要求ステップの処理に基づいて、前記情報検索装置から送信されてくる前記番組情報の受信を制御する受信制御ステップと

を含むことを特徴とするコンピュータが読み取り可能なプログラムが記録されている記録媒体。

【請求項42】 番組の録画を制御する録画制御手段を備える情報処理装置を制御するコンピュータに、

携帯端末装置から送信されてきた、番組を録画予約するための予約情報を分析し、前記番組に関する番組情報を取得するための興味語を抽出する抽出ステップと、

前記抽出ステップの処理により抽出された前記興味語を情報検索装置に送信し、前記興味語に対応する前記番組情報の検索を要求する検索要求ステップと、前記検索要求ステップの処理に基づいて、前記情報検索装置から送信されてくる前記番組情報の受信を制御する受信制御ステップとを実行させるプログラム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、情報検索システム、情報処理装置および方法、情報検索装置および方法、記録媒体、並びにプログラムに関し、特に、例えば、電子メールなどの文書の中から、ユーザの興味があると思われる単語を取得し、その単語に関連する番組情報の推薦を行うようにした情報検索システム、情報処理装置および方法、情報検索装置および方法、記録媒体、並びにプログラムに関する。

【0002】

【従来の技術】

従来、テレビ番組やラジオ番組の推薦を行う方法として、例えば、初期興味登録方式、視聴履歴利用方式、あるいは強調フィルタリング方式などが存在する。

【0003】

いずれの方法も元データとしては、電子番組ガイド（EPG : Electronic Program Guide）やWeb上などの番組情報（番組メタデータ）が用いられており、それらの情報にマッチングすべきユーザの嗜好データの取得方法によって上記3方式に分類されている。

【0004】

初期興味登録方式では、利用開始時、例えば、好きな番組カテゴリ（ドラマやバラエティなど）、好きなジャンル名（ドラマや音楽など）、および好きなタレント名などをユーザに登録してもらい、それ以降、登録された情報をキーワードとして、番組メタデータとのマッチングを取り、推薦すべき番組名を取得する。

【0005】

視聴履歴利用方式では、ユーザが番組を視聴する毎に、その番組に関する番組

メタデータを蓄積し、所定量の視聴履歴（番組メタデータ）が蓄積されると、その視聴履歴を解析して、推薦すべき番組名を取得する。また視聴履歴の代わりに、例えば、ハードディスクドライブに録画する装置などにおいては、ユーザによる録画予約や録画開始といった操作履歴を用いることもできる。この場合、漠然としていた番組ではなく、ユーザの興味をより反映した情報を取得することができる。

【0006】

強調フィルタリング方式では、ユーザの視聴（操作）履歴に対して、他のユーザの視聴履歴との間でマッチングを取り、当該ユーザと視聴履歴の類似する他のユーザの視聴履歴を取得する。そして、視聴履歴が類似する（好みが似ている）他のユーザが視聴した番組で、当該ユーザが未だ視聴していない番組名を取得し、推薦する。

【0007】

以上のような番組推薦方法を用いることによって、ユーザが興味のあると思われる番組を推薦することができる。

【0008】

【発明が解決しようとする課題】

しかしながら、上述した番組推薦方法においては、番組メタデータからユーザの興味を抽出することになり（すなわち、テレビ番組に関する偏った興味を取得することになり）、また、番組メタデータの構造上、一般に通じ易い名称が用いられており、似たような番組が推薦されてしまうという課題があった。

【0009】

すなわち、ユーザの日常の興味を反映することができず、タイムリーに有用な番組などを推薦することができなかつた。

【0010】

また番組推薦時、ユーザは、その番組が推薦されている理由がわからない課題があった。

【0011】

本発明はこのような状況に鑑みてなされたものであり、ユーザが日常使用する

電子メールを分析することにより、ユーザの興味に対応する単語を抽出し、抽出された単語にマッチする番組名を検索して推薦するとともに、その推薦理由を提示できるようにするものである。

【0012】

【課題を解決するための手段】

本発明の第1の情報検索システムは、情報処理装置が、所定の情報を分析し、番組に関する番組情報を取得するための興味語を抽出する抽出手段と、抽出手段により抽出された興味語を情報検索装置に送信し、興味語に対応する番組情報の検索を要求する検索要求手段と、検索要求手段に基づいて、情報検索装置から送信されてくる番組情報を受信する受信手段とを備え、情報検索装置が、番組情報を蓄積する蓄積手段と、情報処理装置から送信されてきた検索要求に基づいて、検索要求に含まれる興味語に関連する番組情報を蓄積手段から検索する検索手段と、検索手段により検索された番組情報を情報処理装置に送信する送信手段とを備えることを特徴とする。

【0013】

情報処理装置の抽出手段には、所定の情報を形態素解析して興味語に分解する形態素解析手段を含むようにすることができる。

【0014】

情報処理装置には、抽出手段により抽出された興味語を用いてデータベースを作成するデータベース構築手段をさらに設けるようにすることができる。

【0015】

情報処理装置には、受信手段により受信された番組情報に基づいて、番組の録画を制御する録画制御手段をさらに設けるようにすることができる。

【0016】

情報処理装置には、受信手段により受信された番組情報の表示を制御する表示制御手段をさらに設けるようにすることができる。

【0017】

情報検索装置の蓄積手段には、番組情報を用いて、番組に関連付けてデータベースを作成するデータベース構築手段を含むようにすることができる。

【0018】

所定の情報は、文書情報、番組に関する嗜好情報、または、番組の視聴履歴に関する情報のうち、少なくとも1つを含むものとすることができます。

【0019】

文書情報は、電子メールであるものとすることができます。

【0020】

番組情報は、番組を録画するための、録画開始時刻、録画終了時刻、およびチャンネル情報を含むものとすることができます。

【0021】

本発明の第1の情報処理装置は、所定の情報を分析し、番組に関する番組情報を取得するための興味語を抽出する抽出手段と、抽出手段により抽出された興味語を情報検索装置に送信し、興味語に対応する番組情報の検索を要求する検索要求手段と、検索要求手段に基づいて、情報検索装置から送信されてくる番組情報を受信する受信手段とを備えることを特徴とする。

【0022】

抽出手段には、所定の情報を形態素解析して興味語に分解する形態素解析手段を含むようにすることができます。

【0023】

抽出手段により抽出された興味語を用いてデータベースを作成するデータベース構築手段をさらに設けるようにすることができます。

【0024】

受信手段により受信された番組情報に基づいて、番組の録画を制御する録画制御手段をさらに設けるようにすることができます。

【0025】

受信手段により受信された番組情報の表示を制御する表示制御手段をさらに設けるようにすることができます。

【0026】

所定の情報は、文書情報、番組に関する嗜好情報、または、番組の視聴履歴に関する情報のうち、少なくとも1つを含むものとすることができます。

【0027】

文書情報は、電子メールであるものとすることができます。

【0028】

番組情報は、番組を録画するための、録画開始時刻、録画終了時刻、およびチャンネル情報を含むものとすることができます。

【0029】

本発明の第1の情報処理方法は、所定の情報を分析し、番組に関する番組情報を取得するための興味語を抽出する抽出ステップと、抽出ステップの処理により抽出された興味語を情報検索装置に送信し、興味語に対応する番組情報の検索を要求する検索要求ステップと、検索要求ステップの処理に基づいて、情報検索装置から送信されてくる番組情報の受信を制御する受信制御ステップとを含むことを特徴とする。

【0030】

本発明の第1の記録媒体に記録されているプログラムは、所定の情報を分析し、番組に関する番組情報を取得するための興味語を抽出する抽出ステップと、抽出ステップの処理により抽出された興味語を情報検索装置に送信し、興味語に対応する番組情報の検索を要求する検索要求ステップと、検索要求ステップの処理に基づいて、情報検索装置から送信されてくる番組情報の受信を制御する受信制御ステップとを含むことを特徴とする。

【0031】

本発明の第1のプログラムは、所定の情報を分析し、番組に関する番組情報を取得するための興味語を抽出する抽出ステップと、抽出ステップの処理により抽出された興味語を情報検索装置に送信し、興味語に対応する番組情報の検索を要求する検索要求ステップと、検索要求ステップの処理に基づいて、情報検索装置から送信されてくる番組情報の受信を制御する受信制御ステップとをコンピュータに実行させる。

【0032】

本発明の第1の情報検索装置は、番組に関する番組情報を蓄積する蓄積手段と、情報処理装置から送信されてきた、番組情報を取得するための興味語を受信す

る受信手段と、受信手段により受信された興味語に関連する番組情報を蓄積手段から検索する検索手段と、検索手段により検索された番組情報を情報処理装置に送信する送信手段とを備えることを特徴とする。

【0033】

興味語は、情報処理装置において所定の情報が形態素解析されて分解された単語であるものとすることができる。

【0034】

番組情報は、番組を録画するための、録画開始時刻、録画終了時刻、およびチャンネル情報を含むものとすることができる。

【0035】

本発明の第1の情報検索方法は、番組に関する番組情報の蓄積を制御する蓄積制御ステップと、情報処理装置から送信されてきた、番組情報を取得するための興味語の受信を制御する受信制御ステップと、受信制御ステップの処理により受信が制御された興味語に関連する番組情報を検索する検索ステップと、検索ステップの処理により検索された番組情報の、情報処理装置への送信を制御する送信制御ステップとを含むことを特徴とする。

【0036】

本発明の第2の記録媒体に記録されているプログラムは、番組に関する番組情報の蓄積を制御する蓄積制御ステップと、情報処理装置から送信されてきた、番組情報を取得するための興味語の受信を制御する受信制御ステップと、受信制御ステップの処理により受信が制御された興味語に関連する番組情報を検索する検索ステップと、検索ステップの処理により検索された番組情報の、情報処理装置への送信を制御する送信制御ステップとを含むことを特徴とする。

【0037】

本発明の第2のプログラムは、番組に関する番組情報の蓄積を制御する蓄積制御ステップと、情報処理装置から送信されてきた、番組情報を取得するための興味語の受信を制御する受信制御ステップと、受信制御ステップの処理により受信が制御された興味語に関連する番組情報を検索する検索ステップと、検索ステップの処理により検索された番組情報の、情報処理装置への送信を制御する送信制

御ステップとをコンピュータに実行させる。

【0038】

本発明の第2の情報検索システムは、携帯端末装置が、番組を録画予約するための予約情報を生成する生成手段と、生成手段により生成された予約情報を情報処理装置に送信する第1の送信手段とを備え、情報処理装置が、携帯端末装置から送信されてきた予約情報を分析し、番組に関する番組情報を取得するための興味語を抽出する抽出手段と、抽出手段により抽出された興味語を情報検索装置に送信し、興味語に対応する番組情報の検索を要求する検索要求手段と、検索要求手段に基づいて、情報検索装置から送信されてくる番組情報を受信する受信手段とを備え、情報検索装置が、番組情報を蓄積する蓄積手段と、情報処理装置から送信されてきた検索要求に基づいて、検索要求に含まれる興味語に関連する番組情報を蓄積手段から検索する検索手段と、検索手段により検索された番組情報を情報処理装置に送信する第2の送信手段とを備えることを特徴とする。

【0039】

情報処理装置の抽出手段には、所定の情報を形態素解析して興味語に分解する形態素解析手段を含むようにすることができる。

【0040】

情報処理装置には、受信手段により受信された番組情報に基づいて、番組の録画を制御する録画制御手段をさらに設けることができる。

【0041】

情報検索装置の蓄積手段には、番組情報を用いて、番組に関連付けてデータベースを作成するデータベース構築手段を含むようにすることができる。

【0042】

予約情報は、番組名、ジャンル名、または、出演者名のうち、少なくとも1つを含むものとすることができます。

【0043】

予約情報は、電子メールであるものとすることができます。

【0044】

番組情報は、番組を録画するための、録画開始時刻、録画終了時刻、およびチ

チャンネル情報を含むものとすることができます。

【0045】

本発明の第2の情報処理装置は、携帯端末装置から送信されてきた、番組を録画予約するための予約情報を分析し、番組に関する番組情報を取得するための興味語を抽出する抽出手段と、抽出手段により抽出された興味語を情報検索装置に送信し、興味語に対応する番組情報の検索を要求する検索要求手段と、検索要求手段に基づいて、情報検索装置から送信されてくる番組情報を受信する受信手段とを備えることを特徴とする。

【0046】

抽出手段には、所定の情報を形態素解析して興味語に分解する形態素解析手段を含むようにすることができます。

【0047】

受信手段により受信された番組情報に基づいて、番組の録画を制御する録画制御手段をさらに設けるようにすることができます。

【0048】

予約情報は、番組名、ジャンル名、または、出演者名のうち、少なくとも1つを含むものとすることができます。

【0049】

予約情報は、電子メールであるものとすることができます。

【0050】

番組情報は、番組を録画するための、録画開始時刻、録画終了時刻、およびチャンネル情報を含むものとすることができます。

【0051】

本発明の第2の情報処理方法は、携帯端末装置から送信されてきた、番組を録画予約するための予約情報を分析し、番組に関する番組情報を取得するための興味語を抽出する抽出ステップと、抽出ステップの処理により抽出された興味語を情報検索装置に送信し、興味語に対応する番組情報の検索を要求する検索要求ステップと、検索要求ステップの処理に基づいて、情報検索装置から送信されてくる番組情報の受信を制御する受信制御ステップとを含むことを特徴とする。

【0052】

本発明の第3の記録媒体に記録されているプログラムは、携帯端末装置から送信されてきた、番組を録画予約するための予約情報を分析し、番組に関する番組情報を取得するための興味語を抽出する抽出ステップと、抽出ステップの処理により抽出された興味語を情報検索装置に送信し、興味語に対応する番組情報の検索を要求する検索要求ステップと、検索要求ステップの処理に基づいて、情報検索装置から送信されてくる番組情報の受信を制御する受信制御ステップとを含むことを特徴とする。

【0053】

本発明の第3のプログラムは、携帯端末装置から送信されてきた、番組を録画予約するための予約情報を分析し、番組に関する番組情報を取得するための興味語を抽出する抽出ステップと、抽出ステップの処理により抽出された興味語を情報検索装置に送信し、興味語に対応する番組情報の検索を要求する検索要求ステップと、検索要求ステップの処理に基づいて、情報検索装置から送信されてくる番組情報の受信を制御する受信制御ステップとをコンピュータに実行させる。

【0054】

本発明の第1の情報検索システムにおいては、情報処理装置で、所定の情報が分析されて番組に関する番組情報を取得するための興味語が抽出され、抽出された興味語に対応する番組情報の検索が要求され、その要求に基づいて情報検索装置から送信されてくる番組情報が受信され、情報検索装置で、情報処理装置から送信されてきた検索要求に基づいて、検索要求に含まれる興味語に関連する番組情報が検索され、検索された番組情報が情報処理装置に送信される。

【0055】

本発明の第1の情報処理装置および方法、並びにプログラムにおいては、所定の情報が分析されて番組に関する番組情報を取得するための興味語が抽出され、抽出された興味語に対応する番組情報の検索が要求され、その要求に基づいて情報検索装置から送信されてくる番組情報が受信される。

【0056】

本発明の情報検索装置および方法、並びに第2のプログラムにおいては、情報

処理装置から送信されてきた番組情報を取得するための興味語が受信され、受信された興味語に関連する番組情報が検索され、検索された番組情報が情報処理装置に送信される。

【0057】

本発明の第2の情報検索システムにおいては、携帯端末装置で、番組を録画予約するための予約情報が生成されて情報処理装置に送信され、情報処理装置で、携帯端末装置から送信されてきた予約情報が分析されて番組に関する番組情報を取得するための興味語が抽出され、抽出された興味語に対応する番組情報の検索が要求され、その要求に基づいて情報検索装置から送信されてくる番組情報が受信され、情報検索装置で、情報処理装置から送信されてきた検索要求に基づいて、検索要求に含まれる興味語に関連する番組情報が検索され、検索された番組情報が情報処理装置に送信される。

【0058】

本発明の第2の情報処理装置および方法、並びに第3のプログラムにおいては、携帯端末装置から送信されてきた、番組を録画予約するための予約情報が分析されて番組に関する番組情報を取得するための興味語が抽出され、抽出された興味語に対応する番組情報の検索が要求され、その要求に基づいて情報検索装置から送信されてくる番組情報が受信される。

【0059】

【発明の実施の形態】

以下、図を参照して、本発明の実施の形態について説明する。

【0060】

図1は、本発明を適用した番組検索システムの構成例を示す図である。この番組検索システムにおいては、インターネットなどのネットワーク5に、利用者端末としてのパーソナルコンピュータ1、ハードディスクドライブ(HDD)レコーダ2、およびデジタル携帯電話機4が接続されているとともに、推薦すべき番組情報(番組メタデータ)を検索するサーバ6が接続されている。パーソナルコンピュータ1は、HDDレコーダ2と、例えば、Ethernet(登録商標)を介して接続されており、HDDレコーダ2には、テレビジョン受像機3が接続されている。す

なわち、パーソナルコンピュータ1、HDDレコーダ2、およびテレビジョン受像機3は、1ユーザ（または1家族）が所有しているものであって、それぞれが近傍に配置されている。

【0061】

パーソナルコンピュータ1は、様々なアプリケーションプログラムの実行が可能な情報処理装置であって、例えば、電子メールの送受信、Webページの閲覧、または文書作成などを行う。パーソナルコンピュータ1はまた、図23のフローチャートを参照して後述するように、電子メールの送受信により取得した文書から、ユーザの興味に対応する単語（以下、適宜、興味語と記述する）を抽出し、興味データのデータベースを作成する。

【0062】

HDDレコーダ2は、テレビ番組を、大容量のハードディスクドライブに録画（記録）するとともに、ユーザからの指示に基づいて、録画したテレビ番組をテレビジョン受像機3に出力し、再生させる。HDDレコーダ2はまた、図25のフローチャートを参照して後述するように、パーソナルコンピュータ1から興味データを取得し、それを、ネットワーク5を介してサーバ6に送信することで、興味データにマッチする番組の推薦を受ける。

【0063】

デジタル携帯電話機4は、図30のフローチャートを参照して後述するように、番組を録画予約するための電子メールを作成し、ネットワーク5を介してパーソナルコンピュータ1またはHDDレコーダ2に送信し、番組の録画予約を実行させる。

【0064】

ネットワーク5は、公衆回線網、移動体無線通信網、ローカルエリアネットワーク、またはインターネットなどのネットワーク、デジタル衛星放送といった、有線または無線のいずれでもよい。

【0065】

図1に示した番組検索システムの例では、利用者端末として、パーソナルコンピュータ1、HDDレコーダ2、テレビジョン受像機3、およびデジタル携帯電話

機4を、それぞれ1つずつしか図示していないが、多数の利用者端末が接続されることは言うまでもない。

【0066】

図2は、パーソナルコンピュータ1にインストールされて実行される、デスクトップマスコット（以下、エージェントと記述する）をデスクトップ上に表示するためのアプリケーションプログラム（以下、エージェントプログラムと記述する）11、電子メールを送受信するためのアプリケーションプログラム（以下、メール（mailer）と記述する）12、および、文書作成または編集するためのワードプロセッサプログラム（以下、ワープロプログラムと記述する）13との関係を説明する図である。

【0067】

エージェントプログラム11は、処理の対象とする文書から、ユーザの興味に応じる単語を抽出し、番組を検索するための興味データ、および、処理の対象とする文書の関連情報を蓄積してデータベースを構築する蓄積部21、処理の対象とする文書に対応する推薦情報をユーザに提示する提示部22、および、エージェント231（図28）の表示等を制御するエージェント制御部23から構成される。

【0068】

なお、蓄積部21および提示部22を、例えばインターネット上の任意のサーバに設置するようにしてもよい。

【0069】

蓄積部21の文書取得部31は、メール12によって送受信された文書やワープロプログラム13によって編集された文書などのうち、自己が未処理の文書を取得して文書属性処理部32および文書内容処理部33に供給する。また、文書取得部31は、ユーザによってHDDレコーダ2に初期登録された嗜好情報（好みのジャンル名や出演者名など）、または視聴履歴を取得し、それらを文書内容処理部33に供給する。

【0070】

なお、以下においては、主に、メール12によって送受信された電子メールの

文書を処理の対象とする場合の例について説明する。

【0071】

文書属性処理部32は、文書取得部31から供給される文書の属性情報を抽出し、その属性情報に基づいて文書をグループ化し、文書内容処理部33および文書特徴データベース作成部34に供給する。電子メールの場合、属性情報としては、文書のヘッダに記述されている情報（対象となっている電子メールを特定するメッセージID、参照している電子メールのメッセージID(References, In-Reply-To)、宛先(To, Cc, Bcc)、あるいは送信元(From)、日付(Date)）、表題(subject）などが抽出される。そして、抽出された属性情報に基づいて、1以上の文書がグループ化される。以下、属性情報に基づいてグループ化された文書群（電子メールグループ）を「話題」と記述する。

【0072】

また、一般にここで言う話題とは、電子メールに限らず、ワープロ、エディタやスケジューラなど、その他のツールやアプリケーションソフトウェアなどから作成されるようなあらゆる文書に関して、ある関係で関連付けられた一連の文書群を指す。

【0073】

文書内容処理部33は、文書属性処理部32でグループ化された文書群（話題）の本文を抽出し、形態素解析を施して、単語（特徴語）に分類する。また、文書内容処理部33は、文書取得部31から供給された嗜好情報または視聴履歴に形態素解析を施して、単語（興味語）に分類する。

【0074】

単語は、品詞（名詞、形容詞、動詞、副詞、接続詞、感動詞、助詞、および助動詞）別に分類される。ただし、広範囲に亘って分布している単語、すなわち、例えば、大多数の文書に含まれていると考えられる単語「こんにちは」、「よろしく」、あるいは「お願いします」等の名詞以外の品詞は関連情報を検索するためのキーワード（以下、検索語とも記述する）には成り得ないので、不要語であるとしてキーワードとする対象から削除される。

【0075】

また、文書内容処理部33は、不要語が削除された各単語の出現頻度および複数の文書に亘る分布状況を求め、グループ化された文書群（話題）毎に、各単語の重み（文書の主旨に関係する程度を示す値、以下、評価値と記述する）を演算する。

【0076】

さらに、文書内容処理部33は、各話題に対し、単語の評価値を要素とする特徴ベクトルを決定する。例えば、各話題に含まれる単語（特徴語）の総数がn個である場合、各話題の特徴ベクトルは、n次元空間のベクトルとして次式（1）のように表現される。

特徴ベクトル

$$= (\text{単語 } 1 \text{ の評価値 } w_1, \text{ 単語 } 2 \text{ の評価値 } w_2, \dots, \text{ 単語 } n \text{ の評価値 } w_n) \\ \dots \quad (1)$$

【0077】

評価値の演算には、例えば文献（Salton,G.:Automatic Text Processing:The Transformation,Analysis, and Retrieval of Information by Computer,Addison-Wesley,1989）に開示されているtf·idf法を用いる。tf·idf法によれば、話題Aに対応するn次元の特徴ベクトルのうち、話題Aに含まれる単語に対応する要素に対しては、評価値として0以外の値が算出され、話題Aに含まれない単語（頻度が0である単語）に対応する要素に対しては、評価値として0が算出される。

【0078】

なお、評価値は、例えば、電子メールの送受信の頻度や回数、電子メールに含まれる単語の品詞の種類（特定の地域や名称を示す固有名詞など）、送受信する相手に応じて修正される。

【0079】

また、本実施の形態においては、話題毎に特徴ベクトルを算出するものとして説明するが、これに限らず、文書毎、または、その他の単位毎（例えば、所定期間（1週間）に蓄積された文書群毎）に特徴ベクトルを算出することも勿論可能である。

【0080】

文書特徴データベース作成部34は、文書属性処理部32によってグループ化された文書群（話題）毎の各文書の属性情報と、文書内容処理部33で算出された話題毎の特徴ベクトル（すなわち、話題内に含まれる単語の評価値）を時系列順にデータベース化するとともに、文書内容処理部33で算出された、嗜好情報または視聴履歴から興味データ（後述）を作成し、それらをデータベース化し、ハードディスクドライブなどよりなる記憶部59に記録する。

【0081】

また、文書特徴データベース作成部34は、単語の評価値などを参照することにより、所定の条件を満たす単語を選択し、関連情報や番組情報を検索するための検索用キーワード（検索語または興味語）として記録する。さらに、文書特徴データベース作成部34は、検索語を関連情報検索部35に供給し、それに対応して関連情報検索部35から供給される関連情報を、検索語に対応付けて記録する。

【0082】

関連情報検索部35は、文書特徴データベース作成部34から供給される検索語に対する関連情報を検索し、検索結果のインデックスを文書特徴データベース作成部34に供給する。検索語に対する関連情報を検索する方法としては、例えばインターネット上の検索エンジンを用いる方法がある。検索エンジンを用いる方法を適用した場合、検索結果として得られるWebページのURL（Uniform Resource Locator）とWebページのタイトルが、関連情報として文書特徴データベース作成部34に供給される。

【0083】

提示部22のイベント管理部41は、メール12がアクティブとされること、メール12が電子メールの送受信を完了したこと、および、入力中の文書のテキストデータ量が所定の閾値を超えたことを検知して、データベース問い合わせ部42に通知する。以下、メール12が電子メールの送受信を完了したこと、または、入力中の文書のテキストデータ量が所定の閾値を超えたことを、イベント発生と記述する。

【0084】

また、イベント管理部41は、内蔵するタイマ41Aを参照することによって時間の経過を監視し、適宜、所定のタイミングから所定の時間が経過した場合、その旨をデータベース問い合わせ部42に通知する。

【0085】

データベース問い合わせ部42は、イベント管理部41からのイベント発生の通知に対応して、イベント発生に対応する文書（例えば、受信した電子メール）を取得し、文書内容処理部33の処理と同様に、その文書に形態素解析を施して単語を抽出し、不要語を除外して各単語の評価値を演算する。これにより、イベント発生に対応する文書の特徴ベクトルが算出される。

【0086】

また、データベース問い合わせ部42は、文書特徴データベース作成部34によって作成されたデータベースを検索し、算出したイベント発生に対応する文書の特徴ベクトルと、データベースに記録されている話題毎の特徴ベクトルとの内積を、両者の類似度として算出する。さらに、データベース問い合わせ部42は、イベント発生に対応する文書に対する類似度が最も高い話題を判定し、その話題に含まれる単語のうち、評価値が所定の条件（詳細については後述する）を満たすものを選択し、選択した単語（重要語）に対応する関連情報（推薦情報）を、イベント管理部41を介し、または直接的に、関連情報提示部43に供給する。

【0087】

さらに、データベース問い合わせ部42は、HDDレコーダ2にログインされたユーザに対応する興味データをデータベースから読み出して、HDDレコーダ2に送信したり、あるいは、イベント発生に対応して、嗜好情報または視聴履歴から作成された興味データをデータベースから読み出して、ネットワーク5を介してサーバ6に送信し、その興味データにマッチする番組情報の検索を要求する。

【0088】

関連情報提示部43は、イベント管理部41を介し、または直接的に、データベース問い合わせ部42から供給される関連情報（推薦情報）を表示部58（デ

スクトップ) 上に表示させる。すなわち、イベント管理部41がイベント発生を検知する毎、提示部22による関連情報の提示が更新される。

【0089】

なお、蓄積部21によるデータベースの更新は、所定のタイミングにおいて実行される。データベースの更新処理は、図21のフローチャートを参照して後述する。また、蓄積部21によるデータベースの更新時には、記憶部59に記録した特徴ベクトルが、例えば、電子メールの送受信の頻度や回数、電子メールに含まれる単語の品詞の種類（特定の地域や名称を示す固有名詞など）に応じて修正される。

【0090】

図示は省略するが、HDDレコーダ2にインストールされて実行されるエージェントプログラムも、上述した図2のエージェントプログラム11とほぼ同様の機能を有する。なお、HDDレコーダ2が、パーソナルコンピュータ1の蓄積部21を使用（共有）することで、その部分を省略するようにしてもよい。

【0091】

図3は、エージェントプログラム11乃至ワープロプログラム13がインストールされて実行されるパーソナルコンピュータ1の構成例を示している。なお、当然ながら、本発明はパーソナルコンピュータの他、ホームサーバシステム、ゲーム機器、カーナビゲーションシステム、PDA (Personal Digital Assistant) 等の情報電子機器において利用できる。

【0092】

このパーソナルコンピュータ1は、CPU(Central Processing Unit)51を内蔵している。CPU51には、バス54を介して入出力インターフェース55が接続されている。入出力インターフェース55には、キーボード、マウスなどの入力デバイスよりなる入力部56、処理結果としての例えば音声信号を出力する出力部57、処理結果としての画像を表示するディスプレイなどよりなる表示部58、プログラムや構築されたデータベースなどを格納するハードディスクドライブなどよりなる記憶部59、インターネットに代表されるネットワークを介してデータを通信するLAN (Local Area Network) カードなどよりなる通信部60、および、

磁気ディスク62、光ディスク63、光磁気ディスク64、または半導体メモリ65などの記録媒体に対してデータを読み書きするドライブ61が接続されている。バス54には、ROM(Read Only Memory)52およびRAM(Random Access Memory)53が接続されている。

【0093】

本発明のエージェントプログラム11は、磁気ディスク62乃至半導体メモリ65に格納された状態でパーソナルコンピュータに供給され、ドライブ61によって読み出されて、または通信部60がネットワークを介して取得して、記憶部59に内蔵されるハードディスクドライブにインストールされている。記憶部59にインストールされているエージェントプログラム11は、入力部56に入力されるユーザからのコマンドに対応するCPU51の指令によって、記憶部59からRAM53にロードされて実行される。なお、パーソナルコンピュータ1の起動時において自動的にエージェントプログラム11が実行されるように設定することも可能である。

【0094】

また記憶部59に内蔵されるハードディスクドライブには、エージェントプログラム11の他、メール12、およびワープロプログラム13、WWW(World Wide Web)ブラウザなどのアプリケーションプログラムもインストールされており、エージェントプログラム11と同様に、入力部56に入力されるユーザからの起動コマンドに対応するCPU51の指令によって、記憶部59からRAM53にロードされて実行される。

【0095】

図4は、HDDレコーダ2の構成例を示すブロック図である。このHDDレコーダ2は、膨大な記憶容量を備えるハードディスクドライブ(HDD)78に数多くの映像を録画することができるとともに、ユーザの意図を的確に把握して録画映像の記録管理(視聴履歴や操作履歴など)を反映することができる。なお、HDDレコーダ2は、AV機器として実装することができ、例えば、セットトップボックス(STB)のようなテレビジョン受像機と一体で構成することができる。

【0096】

CPU 7 1は、HDDレコーダ2全体の動作を制御するメインコントローラであり、入力部7 6から供給された入力信号に基づいて、チューナ7 9、復調部8 0、デコーダ8 1、およびHDD7 8などを制御することにより、番組の録画または再生を実行させる。

【0097】

RAM7 3は、CPU7 1の実行プログラムをロードしたり、実行プログラムの作業データを書き込むために使用される、書き込み可能な揮発性メモリである。ROM7 2は、HDDレコーダ2の電源オン時に実行する自己診断・初期化プログラムや、ハードウェア操作用の制御コードなどを格納する読み出し専用メモリである。

【0098】

入力部7 6は、例えば、リモートコマンダ、ボタン、スイッチ、あるいはキーボードなどにより構成され、操作に対応する入力信号を、入出力インターフェース7 5およびバス7 4を介してCPU7 1に出力する。

【0099】

通信部7 7は、ネットワーク5を介して、サーバ6と通信することにより、推薦された番組メタデータを受信するとともに、パーソナルコンピュータ1と通信することにより、所定データ（例えば、興味データ）を送受信することができる。通信部7 7に入力されたデータは、適宜、入出力インターフェース7 5を介してHDD7 8に記録される。

【0100】

HDD7 8は、プログラムやデータなどを所定フォーマットのファイル形式で蓄積することができる、ランダムアクセスが可能な記憶装置であり、膨大な記憶容量を備えている。HDD7 8は、入出力インターフェース7 5を介してバス7 4に接続されており、デコーダ8 1、または通信部7 7から、放送番組、およびEPGデータなどのデータ放送用のデータの入力を受け、これらの情報を記録するとともに、必要に応じて、記録されたデータを出力する。

【0101】

図示せぬアンテナで受信された放送波は、チューナ7 9に供給される。放送波は、所定のフォーマットに基づいており、例えば、EPGデータを含んでいても良

い。放送波は、衛星放送波、地上波、有線、または無線の区別を特に問わない。

【0102】

チューナ79は、CPU71の制御に基づいて、所定チャンネルの放送波のチューニングすなわち選局を行い、受信データを復調部80に出力する。なお、送信されてくる放送波がアナログであるか、あるいは、デジタルであるかに応じて、チューナ79の構成を適宜変更または拡張することができる。復調部80は、デジタル変調されている受信データを復調し、デコーダ81に出力する。

【0103】

例えば、デジタル衛星放送の場合、チューナ79により受信され、復調部80により復調されたデジタルデータは、MPEG2 (Moving Picture Experts Group 2) 方式で圧縮されたAVデータ、および、データ放送用のデータが多重化されているトランスポートストリームである。前者のAVデータは、放送番組本体を構成する映像データおよび音声データであり、後者のデータ放送用のデータは、この放送番組本体に付随するデータ（例えば、EPGデータ）を含むものである。

【0104】

デコーダ81は、復調部80より供給されたトランスポートストリームを、MPEG方式で圧縮されたAVデータとデータ放送用のデータ（例えば、EPGデータ）に分離する。分離されたデータ放送用のデータは、バス74および入出力インターフェース75を介してHDD78に供給され、記録される。

【0105】

受信した番組をそのまま出力することが指示されている場合、デコーダ81は、AVデータを、圧縮映像データと圧縮音声データにさらに分離する。分離された音声データは、デコードされた後、ミキサ83を介してテレビジョン受像機3のスピーカに出力される。また、分離された映像データは、伸張された後、コンポーザ82を介してテレビジョン受像機3のモニタに出力される。

【0106】

受信した番組をHDD78に録画することが指示されている場合、デコーダ81は、分離する前のAVデータを、バス74および入出力インターフェース75を介してHDD78に出力する。また、HDD78に録画されている番組の再生が指示され

ている場合、デコーダ81は、入出力インターフェース75およびバス74を介して、HDD78からAVデータの入力を受け、圧縮映像データと圧縮音声データに分離し、それぞれコンポーザ82もしくはミキサ83に出力する。

【0107】

コンポーザ82は、必要に応じて、デコーダ81より入力された映像データと、GUI(Graphical User Interface)画面とを合成して、テレビジョン受像機3のモニタに出力する。

【0108】

図5は、サーバ6にインストールされて実行されるサーバプログラム101の機能を説明する図である。

【0109】

サーバプログラム101は、処理の対象とする、EPGデータなどの番組メタデータを分析し、推薦番組用データベースを構築する蓄積部111、および、蓄積部111に蓄積されている推薦番組用データベースの中から、ユーザの興味データにマッチする番組情報を検索する検索部112から構成される。

【0110】

蓄積部111の番組メタデータ取得部121は、EPGデータなどの番組メタデータのうち、自己が未処理の番組メタデータを、図示せぬEPGデータ提供装置から取得してデータ内容処理部122に供給する。

【0111】

データ内容処理部122は、番組メタデータ取得部121から供給される番組メタデータに形態素解析を施し、番組情報（番組名、ジャンル名、放送局、時間帯、出演者、およびキーワードなど）を抽出する。抽出された番組情報は、データベース作成部123に供給される。

【0112】

データベース作成部123は、データ内容処理部122によって抽出された番組情報を番組毎にデータベース化して、ハードディスクドライブなどよりなる記憶部147（図6）に記録する。

【0113】

検索部112のイベント管理部131は、ネットワーク5を介して利用者端末装置（パーソナルコンピュータ1またはHDDレコーダ2）から興味データの入力を検知して、データベース問い合わせ部132に通知する。以下、興味データの入力を検知したことを、検索要求と記述する。また、イベント管理部131は、内蔵するタイマ131Aを参照することによって時間の経過を監視し、適宜、所定のタイミングから所定の時間が経過した場合、その旨をデータベース問い合わせ部132に通知する。

【0114】

データベース問い合わせ部132は、イベント管理部131からの検索要求の通知に基づいて、検索要求に対応する興味データを取得する。データベース問い合わせ部132は、検索エンジンを用いて、データベース作成部123によって作成された推薦番組用データベースを検索し、取得した興味データにマッチする番組情報を推薦用番組として選択する。

【0115】

また、データベース問い合わせ部132は、利用者端末装置で初期登録された嗜好情報（好みのジャンル名や出演者名など）、または視聴履歴などを取得し、データ内容処理部122の処理と同様に、嗜好情報や視聴履歴に形態素解析を施して、興味データを抽出し、データベース作成部123の推薦番組用データベースを検索し、抽出した興味データにマッチする番組情報を推薦用番組として選択する。

【0116】

以上のようにして選択された推薦用番組は、イベント管理部131を介し、または直接的に、番組情報出力部133に供給される。

【0117】

番組情報出力部133は、イベント管理部131を介し、または直接的に、データベース問い合わせ部132から供給される推薦用番組（番組情報）を、ネットワーク5を介して利用者端末装置（パーソナルコンピュータ1またはHDDレコーダ2）に検索結果として出力する。

【0118】

なお、蓄積部111による推薦番組用データベースの更新は、EPGデータが更新される毎、あるいは、所定のタイミングにおいて実行される。

【0119】

図6は、サーバプログラム101がインストールされて実行されるサーバ6の構成例を示している。同図に示されるCPU141乃至半導体メモリ153は、図3に示したパーソナルコンピュータ1のCPU51乃至入力部56並びに記憶部59乃至半導体メモリ65と基本的に同様の構成を有するものであるため、その説明は省略する。

【0120】

次に、パーソナルコンピュータ1のエージェントプログラム11によるデータベース作成処理について、図7のフローチャートを参照して説明する。このデータベース作成処理は、エージェントプログラム11が実行する処理のうちの1つであり、エージェントプログラム11が起動された状態において、データベースが未だ作成されていないときに開始される。

【0121】

ステップS1において、文書取得部31は、データベース作成の素として分析する文書（例えば、エージェントプログラム11が実行される以前に送受信された電子メール、以下、分析対象電子メールと記述する）を、記憶部59に内蔵されるハードディスクドライブから選択的に取得して文書属性処理部32および文書内容処理部33に供給する。

【0122】

ステップS1の処理、すなわち、分析対象電子メール選択処理の詳細について、図8を参照して説明する。

【0123】

ステップS21において、文書取得部31は、ユーザが送信した電子メールが保存されている送信フォルダを参照し、直近の所定期間（例えば、最近の一週間）に送信した電子メールの数が所定数（例えば、100通）以上存在するか否かを判定する。直近の所定期間に送信した電子メールの数が所定数以上存在すると判定された場合、処理はステップS22に進む。ステップS22において、文書

取得部31は、日時条件およびアドレス属性条件を設定する。

【0124】

ステップS22の処理、すなわち、日時条件およびアドレス属性条件を設定処理の詳細について、図9を参照して説明する。ステップS31において、文書取得部31は、送信フォルダに存在する電子メールの数が所定数（例えば、10000通）以上であるか否かを判定する。

【0125】

ステップS31において、送信フォルダに存在する電子メールの数が所定数以上であると判定された場合、処理はステップS32に進む。ステップS32において、文書取得部31は、分析対象電子メールを選択するための日時条件を「1年以前を除去」に設定する。ステップS33において、文書取得部31は、分析対象電子メールを選択するためのアドレス属性条件を「"To"以外を除去」に設定する。また、文書取得部31は、アドレス条件（アドレスリスト）を抽出する対象を送信フォルダに設定する。

【0126】

反対に、ステップS31において、送信フォルダに存在する電子メールの数が所定数よりも少ないと判定された場合、処理はステップS34に進む。ステップS34において、文書取得部31は、日時条件を「3年以前を除去」に設定する。ステップS35において、文書取得部31は、アドレス属性条件を「"To, Cc"以外を除去」に設定する。また、文書取得部31は、アドレス条件を抽出する対象を送信フォルダおよび受信フォルダに設定する。

【0127】

以上のような日時条件およびアドレス属性条件設定処理により、送信した電子メールの数に対応して、分析対象電子メールの日時条件とアドレス属性条件が設定された後、処理は、図8のステップS23にリターンする。

【0128】

なお、日時条件およびアドレス属性条件設定処理は、上述した2種類の選択だけでなく、例えば、送信フォルダのメール数に応じていくつかの区間を設け、それに応じて、日時条件を任意の年数で細かく区切ったり、受信簿に対するアドレ

ス属性条件にさらにfrom, reply to等を加えた選択肢を増やすなどしてもよい。

【0129】

ステップS23において、文書取得部31は、送信フォルダ（または受信フォルダ）に存在する電子メールを、ステップS22で設定した日時条件およびアドレス属性条件に基づいてフィルタリングすることにより、電子メールの数を絞り込む。ステップS24において、文書取得部31は、ステップS23でフィルタリングされた各電子メールの宛先（または送信元）をリスト化するとともに、各宛先の出現回数をカウントし、出現回数が多い上位n個のアドレスを判定して、アドレス条件を「上位n個のアドレスから送受信された電子メールを抽出」に設定する。

【0130】

ステップS25において、文書取得部31は、全ての電子メール、すなわち、送信フォルダ、受信フォルダ、およびその他のフォルダの存在する電子メールのうち、ステップS22で設定した日時条件およびステップS24で設定したアドレス条件に基づいてフィルタリングすることにより、分析対象電子メールを選択する。

【0131】

なお、ステップS21において、ユーザが送信した電子メールが保存されている送信フォルダを参照し、直近の所定期間に送信した電子メールの数が所定数よりも少ないと判定された場合、処理はステップS26に進む。ステップS26において、文書取得部31は、ユーザが送信した電子メールが保存されている受信フォルダを参照し、直近の所定期間（例えば、最近の一週間）に受信した電子メールの数が所定数（例えば、100通）以上存在するか否かを判定する。直近の所定期間に受信した電子メールの数が所定数以上存在すると判定された場合、処理はステップS22に進み、それ以降の処理が繰り返される。

【0132】

反対に、ステップS26において、直近の所定期間に受信した電子メールの数が所定数よりも少ないと判定された場合、この段階でデータベース作成処理は終了される。

【0133】

以上のように分析対象電子メールが選択された後、処理は、図7のステップS2にリターンする。

【0134】

ステップS2において、文書属性処理部32は、ステップS1の処理で文書取得部31から供給された分析対象電子メールから属性情報（メッセージID等のヘッダ情報）を抽出し、その属性情報に基づき、分析対象電子メールを話題毎に分類して（すなわち、話題毎にグループ化して）、話題毎に話題ファイルを生成して文書内容処理部33および文書特徴データベース作成部34に供給する。

【0135】

図10は、ステップS2において作成される話題ファイル161の一例を示している。話題ファイル161は、各話題ファイルを識別するためのトピックスID162、当該話題に属する最古の電子メールの通信時間を示す日時情報163、当該最古の電子メールの題名などを示すサブジェクト情報164、当該話題に属する電子メールの送信元または宛先の電子メールアドレスからなるメンバー情報165、当該話題に属する各電子メールを特定するメールメッセージID166、当該話題に属する電子メールの本文に含まれる単語から構成される単語ベクトル167、当該話題に属する電子メールの本文を連結した連結本文168、およびいずれかの話題に含まれる全ての単語の評価値から成る特徴ベクトル169から構成される。

【0136】

トピックスID162として、例えば、当該話題に属する最古の電子メールの通信時間を用いるようにしてもよい。

【0137】

なお、連結本文168は、当該話題に属する電子メールのうち、送信フォルダに存在する電子メールの本文を連結した後、所定の文字列（例えば”soshin-shu ryo”）を挿入して、受信フォルダやその他のフォルダに存在する電子メールの本文を連結するようにする。

【0138】

図11は、単語ベクトル167を構成する複数の単語170に含まれる要素を示している。すなわち、単語170には、当該単語自身の文字列171、当該単語の品詞（名詞の種類）172、当該話題における当該単語の頻度173、および当該話題における当該単語の評価値174を記録するための構成を有している。なお、単語170の各要素の中身は、ステップS2の処理段階では生成されず、以降の処理において生成される。

【0139】

また、特徴ベクトル169も、ステップS2の処理段階では生成されず、以降の処理において生成される。

【0140】

図7に戻る。ステップS3において、文書属性処理部32は、ステップS2で生成した話題を選抜する。ステップS3の処理、すなわち第1次話題選抜処理について、図12のフローチャートを参照して説明する。

【0141】

ステップS41において、文書属性処理部32は、ステップS2で生成した話題の数が所定数以上存在するか否かを判定する。生成した話題の数が所定数以上存在すると判定された場合、処理はステップS42に進む。ステップS42において、文書属性処理部32は、生成した話題を選抜するための構成メール数条件を「a（例えば4）通以下を削除」に設定する。

【0142】

反対に、ステップS41において、生成した話題の数が所定数よりも少ないと判定された場合、処理はステップS43に進む。ステップS43において、文書属性処理部32は、生成した話題を選抜するための構成メール数条件を「b（aよりも小さい数、例えば2）通以下を削除」に設定する。

【0143】

ステップS44において、文書属性処理部32は、上段の処理で設定した構成メール数条件に基づき、ステップS2で生成した話題をフィルタリングする。すなわち、例えば、上段の処理で構成メール数条件を「a通（例えば4通）以下を削除」に設定した場合、4通以下の電子メールから構成される話題を削除し、5

通以上の電子メールから構成される話題だけを選抜する。

【0144】

さらに、直近の所定期間（例えば、最近の一週間）に通信した電子メールを含まない話題を削除するようにしてもよい。

【0145】

このようにして第1次話題選抜処理を実行した後、処理は、図7のステップS4にリターンする。

【0146】

なお、第1次話題選抜処理における構成メール数条件の設定は、上述した2種類の選択だけでなく、例えば、話題の数に応じていくつかの区間を設けて、その区間ごとに構成メール数条件を決定するようにしてもよい。

【0147】

ステップS4において、文書内容処理部33は、選抜された各話題に対応する話題ファイル161の連結本文168に形態素解析を実行する。ステップS4における形態素解析処理の詳細について、図13のフローチャートを参照して説明する。

【0148】

ステップS51において、文書内容処理部33は、選抜された各話題のうち、形態素解析を施していないものが存在するか否かを判定する。形態素解析を施していないものが存在すると判定された場合、処理はステップS52に進む。ステップS52において、文書内容処理部33は、形態素解析を施していない話題を1つ選択し、対応する話題ファイル161の連結本文168を読み出して形態素解析を施し、連結本文168に含まれる単語を抽出する。

【0149】

このように、話題ファイル161の連結本文168に対して形態素解析を施す処理は、話題ファイル161を構成する電子メールの各本文に対して形態素解析を施す処理に比較して、処理する文章は長くなるが処理回数が1回で済むので、処理に要するリソースの浪費を抑止することができる。

【0150】

ステップS53において、文書内容処理部33は、ステップS52で抽出した単語のうち、品詞が名詞（一般名詞、サ変接続名詞、地名、人名、興味がある用語を含む）であるものを抽出する。ステップS54において、文書内容処理部33は、抽出した名詞である単語を並べ、当該話題に対応する単語ベクトル167を生成する。

【0151】

ステップS55において、文書内容処理部33は、話題単語テーブル181（図14）にステップS54で生成した単語ベクトル167に対応する記録を追加するとともに、ステップS54で生成した単語ベクトル167を構成する単語の記録を、話題評価値テーブル193を含む単語インデックステーブル191（図15）に追加する。なお、話題単語テーブル181、単語インデックステーブル191、および話題評価値テーブル193は、いずれもハッシュテーブル（Hash table）である。

【0152】

図14は、話題単語テーブル181の構成例を示している。話題単語テーブル181は、各話題に対するトピックスID162と、それに対応する単語ベクトル167が記録されており、トピックスID162を入力として、対応する単語ベクトル167を出力する。

【0153】

図15は、単語インデックステーブル191の構成例を示している。単語インデックステーブル191は、各単語ベクトル167を構成する単語名192と、それに対応する話題評価値テーブル193の組が複数記録されており、単語名192を入力として、話題評価値テーブル193を出力する。

【0154】

図16は、話題評価値テーブル193の構成例を示している。話題評価値テーブル193は、単語名192に対応する単語が含まれる話題のトピックスID201と、当該話題における当該単語の評価値202が記録されており、トピックスID201を入力として、当該話題における当該単語の評価値202を出力する。

【0155】

このような構成の話題単語テーブル181乃至話題評価値テーブル193を生成することにより、トピックスID162および単語名192のどちらか一方を入力としても、対応する他方を容易に検索することが可能となる。

【0156】

この後、処理はステップS51に戻り、以降の処理が繰り返される。その後、ステップS.51において、選抜された各話題のうち、形態素解析を施していないものが存在しないと判定された場合、形態素解析処理は終了され、処理は、図7のステップS5にリターンする。

【0157】

ステップS5において、文書内容処理部33は、以降における処理を軽減するために、これまでの処理で抽出された単語、すなわち、各話題にそれぞれ対応する単語ベクトルに含まれる単語のうち、話題の内容に関連が薄いと考えられる単語、あいさつなどの日常的な単語等（以下、不要語と記述する）を除去する。

【0158】

ステップS5における不要語削除処理について、図17のフローチャートを参照して説明する。ステップS61において、文書内容処理部33は、単語ベクトルが小さい話題、すなわち、対応する単語ベクトルを構成する単語の数が所定数（例えば、5個）以下である話題を除去する。

【0159】

ステップS62において、文書内容処理部33は、ステップS4の処理で生成した単語インデックステーブル191に記録されている単語のうち、以降の処理の対象としていない単語が存在するか否かを判定する。処理対象としていない単語が存在すると判定された場合、処理はステップS63に進む。ステップS63において、文書内容処理部33は、単語インデックステーブル191に記録されている、処理対象としていない単語のうちの1つを処理対象の単語に選択する。

【0160】

ステップS64において、文書内容処理部33は、処理対象の単語を入力として、単語インデックステーブル191を参照することにより、対応する話題評価

テーブル193を取得し、取得した話題評価テーブル193に記録されているトピックスID201の数をカウントすることによって、処理対象の単語を含む話題の数を取得する。

【0161】

ステップS65において、文書内容処理部33は、処理対象の単語を含む話題の数が所定数以上であるか否かを判定する。処理対象の単語を含む話題の数が所定数以上であると判定された場合、処理はステップS66に進む。ステップS66において、文書内容処理部33は、処理対象の単語を、不要語ベクトル（不要語を構成要素とする）に追加する。これにより、多数の話題に共通して含まれると考えられるあいさつなどの日常的な単語が不要語ベクトルに追加される。

【0162】

ステップS67において、文書内容処理部33は、不要語である処理対象の単語に対応する記録を削除するため、各話題にそれぞれ対応する話題ファイル161、話題単語テーブル181、単語インデックステーブル191、および話題評価値テーブル193を更新する。この後、処理はステップS62に戻り、以降の処理が繰り返される。

【0163】

なお、ステップS65において、処理対象の単語を含む話題の数が所定数よりも小さいと判定された場合にも、ステップS66およびS67はスキップされて、処理は、ステップS62にリターンする。

【0164】

その後、ステップS62において、ステップS4の処理で生成した単語インデックステーブル191に記録されている単語のうち、以降の処理の対象としていない単語が存在しないと判定された場合、処理はステップS68に進む。ステップS68において、文書内容処理部33は、再びステップS61の処理と同様に、単語ベクトルが小さい話題、すなわち、対応する単語ベクトル167を構成する単語の数が所定数（例えば、5個）以下である話題を除去する。これにより、日常的な単語ばかりで構成されているとみなされる話題が除去される。この段階で、話題は特徴的な単語から構成される単語ベクトル167によって象徴される

ことになる。処理は、図7のステップS6にリターンする。

【0165】

ステップS6において、文書内容処理部33は、不要語が削除された各単語ベクトル67を構成する全ての単語について、その出現頻度および複数の文書に亘る分布状況を求め、各話題における評価値を演算する。評価値の演算には、例えばtf·idf法を用いる。ステップS7において、文書特徴データベース作成部34は、ステップS6で演算した各単語に対する評価値を、次の条件に基づいて修正する。

【0166】

例えば、送信した電子メールに含まれる単語の評価値がより大きくなるように修正を行う。送信した電子メールに含まれる単語を特定するためには、ステップS2の処理で生成した各話題に対応する話題ファイル161の連結本文168に挿入した、所定の文字列（例えば”soshin-shuryo”）を検出し、当該所定の文字列以前の単語を、送信した電子メールに含まれる単語として特定すればよい。

【0167】

また例えば、属する電子メールの数が多い話題に含まれる単語の評価値が、属する電子メールの数に対応して大きくなるように修正を行う。例えば、属する電子メールの数をmとした場合、修正前の評価値に対し、1次関数値 $a \cdot m$ （aは定数）、対数関数値 $\log(m)$ などの単調増加関数値を乗算する。この修正は、電子メールのような時間的に継続するやりとりでは、以前の文書に登場した単語が、次の文書では指示代名詞によって置換されることが多いので、話題に属する電子メールの数が多くなるほど、単語の評価値が相対的に小さくなってしまう傾向にあることを考慮したものである。

【0168】

さらに例えば、通信頻度が高い相手と通信した電子メールに含まれる単語、および特定名詞（定義した興味語、一般名、地名、組織名など）などの評価値がより大きくなるように修正を行う。なお、特定名詞に対する評価値の修正方法については、特願2001-379511号として提案した発明を適用することができる。

【0169】

ステップS8において、文書特徴データベース作成部34は、ステップS6で演算され、ステップS7で修正された各単語に対する評価値を、話題ファイル161および話題単語テーブル181の単語ベクトル167、並びに単語インデックステーブル191の中の話題評価値テーブル193に記録する。これにより、各単語ベクトル167を構成する単語170の全ての要素が決定されることになる。また、文書特徴データベース作成部34は、各話題にそれぞれ対応する特徴ベクトル169を確定して記録する。さらに、文書特徴データベース作成部134は、各単語ベクトル167について、構成する単語を評価値が大きい順に並べ替える。

【0170】

ステップS9において、文書特徴データベース作成部34は、この段階で残っている話題を再び選抜する。ステップS9の処理、すなわち第2次話題選抜処理について、図18のフローチャートを参照して説明する。なお、この第2次話題選抜処理は、各話題に対して実行される。

【0171】

ステップS71において、文書特徴データベース作成部34は、話題に対応する単語ベクトル167を構成する単語のうち、評価値が最大のもの（あるいは、上位の2、3語）を検出する。ステップS72において、文書特徴データベース作成部34は、ステップS71で検出した単語の評価値が所定値以上であるか否かを判定する。検出した単語の評価値が所定値以上であると判定された場合、処理はステップS73に進む。

【0172】

ステップS73において、文書特徴データベース作成部34は、評価値が所定値以上の当該単語を推薦話題候補ベクトルに追加する。ステップS72において、ステップS71で検出した単語の評価値が所定値よりも小さいと判定された場合、処理はステップS74に進み、文書特徴データベース作成部34は、当該話題を削除する。すなわち、評価値が所定値より小さい単語は興味が少ないものと判断されて検索対象から除外される。

【0173】

ステップS73またはステップS74の処理の後、すなわち、当該話題に対する第2次話題選抜処理が終了された後、次の話題に対する第2次話題選抜処理が開始される。全ての話題について第2次話題選抜処理が終了されると、処理は、図7のステップS10にリターンする。

【0174】

ステップS10において、文書特徴データベース作成部34は、ステップS9の処理で推薦話題候補ベクトルに追加された話題に基づいて、推薦話題を確定する。ステップS10における推薦話題確定処理について、図1.9のフローチャートを参照して説明する。

【0175】

ステップS81において、文書特徴データベース作成部34は、ステップS9の処理で推薦話題候補ベクトルに追加された要素（各単語ベクトル167）から、構成する単語の評価値の最大値に注目し、評価値の最大値が大きい順に所定数（例えば、200）だけ単語ベクトル167を検出し、それぞれに対応する所定数の話題を取得する。

【0176】

ステップS82において、文書特徴データベース作成部34は、ステップS81の処理で取得された話題が検索条件に合致するか否かを判定し、検索条件に合致すると判定した場合、ステップS83に進む。ここで検索条件とは、例えば、特定期間の話題であるか、特定の相手とやりとりした話題であるか、特定の単語を含む話題であるか、視聴履歴から抽出された単語（例えば、番組名、ジャンル名、あるいは出演者名）を含む話題であるか、または、初期登録された嗜好情報（好みのジャンル名や出演者名）を含む話題であるなどである。

【0177】

ステップS83において、文書特徴データベース作成部34は、検索条件に合致した話題を推薦話題ベクトルに追加する。ステップS82において、取得された話題が検索条件に合致しないと判定された場合、ステップS83の処理はスキップされる。その後、処理は、図7のステップS11にリターンする。

【0178】

ステップS11において、文書特徴データベース作成部34は、ステップS10の処理で確定された推薦話題ベクトルに基づいて、データベースを作成する。

【0179】

具体的には、推薦話題ベクトルに追加されている話題の話題ファイル161(図10)をフィルタリングし、後述する番組情報の検索処理において必要となるトピックスID162、日時情報163、サブジェクト情報164、メンバー情報165、および単語ベクトル167(興味語ベクトル212)を抽出して、図20に示されるような興味データ211を作成する。そして、文書特徴データベース作成部34は、興味データ211を、当該ユーザのユーザID、メールアカウント、ログインアカウント、またはパスワードなどに対応付けてデータベースを作成し、記憶部59に記録する。ただし、ステップS11の処理は、ステップS10までの一連の処理に継続して実行される場合と、一連の処理に継続せず、所定のタイミングにおいて実行される場合がある。

【0180】

なお、興味データ211は、番組情報を検索するためのキーワードとなるため、単語ベクトル167を、新たに興味語ベクトル212として定義する。また単語ベクトル167を構成する単語(興味語)のうち、評価値が最大のもの(あるいは、上位の2、3語)のみを興味語ベクトル212として用いるようにしてもよい。また、興味データ211には、例えば、当該話題の電子メールの重要語を含む1文書を保持させるようにしてもよい。

【0181】

ところで、テレビ番組のEPGデータでよく使用される単語、例えば、「ミュージック」、「ニュース」、または、出演者のニックネームなどの未定義語(表記ゆれや短縮語を含む)は、例えば、“名詞-テレビ語”といったように、単なる一般名詞とは異なるように、予め定義しておく。そして、形態素解析の結果、“名詞-テレビ語”に分類された単語は、図7のステップS6における評価値の演算時に、所定の重みがかけられ、“名詞-テレビ語”的な単語の評価値がより大きくなるように修正される。このようにしておくことにより、電子メールに含まれ

る単語の中で“名詞－テレビ語”の単語が、興味語ベクトル212を構成する単語に含まれる（すなわち、興味語になる）可能性が高まる。

【0182】

さらに、複数のユーザが、1つのメール12を共用している場合、メールアカウント毎に、電子メールの形態素解析が施され、ユーザ毎に興味データ211が作成される。データベースには、メールアカウントをキーにして、ユーザ毎の興味データ211が記録される。

【0183】

以上のデータベース作成処理が実行されることにより、送受信した電子メールの文書から抽出された興味語で構成される興味データ211がデータベース内に蓄積されることになる。

【0184】

また、データベース作成処理をユーザが強制的に中断することができるようするため、中断要求があった場合、中断された時点で処理済みの文書を記録し、再開要求があった場合、未処理の文書から処理を再開するようにしてもよい。

【0185】

なお、データベース作成処理は、エージェントプログラム11が実行された場合に開始されるものとしたが、任意のタイミングで開始させることも可能である。さらに、このようにして作成されたデータベースは、所定の条件が満たされたときに更新される。

【0186】

次に、蓄積部21によるデータベースの更新のタイミングについて説明する。データベースは、上述したデータベース作成処理によって作成されるが、次のような第1乃至第3の状況になった場合、データベースが更新される。

【0187】

すなわち、第1の状況として、データベースが作成または更新されてから所定の期間が経過している場合、データベース内の関連情報が古くなってしまうので更新が行われる。

【0188】

第2の状況として、データベースに蓄積されている関連情報のうちの所定の割合が提示済みとなった場合、データベース内の同じ関連情報が繰り返し提示されたり、提示する関連情報が不足したりしてしまって更新が行われる。

【0189】

第3の状況として、特徴抽出に用いた文書が電子メールである場合、電子メールの送受信が繰り返されていると、その文書の内容が変化するので更新が行われる。

【0190】

なお、データベースの更新が必要である状況になった場合（例えば、イベント管理部41がタイマ41Aを監視し、所定の期間が経過したとき）、ユーザに対して更新を指示するように促すこともできるし、ユーザに対する更新指示の促しを実行することなく、自動的にデータベースを実行するように設定することも可能である。また、ユーザが指定する任意のタイミングで更新することも勿論可能である。

【0191】

これら第1乃至第3の状況を考慮したデータベース更新処理について、図21のフローチャートを参照して説明する。このデータベース更新処理は、エージェントプログラム11が実行する処理のうちの1つであり、エージェントプログラム11の起動とともに開始され、エージェントプログラム11が終了されるまで繰り返し実行される。なお、この処理が開始される以前において、既に上述したデータベース作成処理が実行されており、データベースが存在するものとする。

【0192】

ステップS91において、エージェントプログラム11の蓄積部21は、作成済みのデータベースの更新が必要であるか否かを判定し、更新が必要であると判定されるまで待機する。この判断基準は、例えば図22に示すようなユーザインターフェースの画面を用いて予めユーザが設定するものとする。図22の例では、4つの条件が示されており、ユーザによって左端の口印（チェックボックス）がチェックされた場合、対応する条件が有効となる。なお、1番目の条件では回数が設定可能とされており、3番目の条件では日数が設定可能とされている。

【0193】

ステップS91において、更新が必要であると判定された場合、処理は、ステップS92に進む。ステップS92において、蓄積部21は、データベースを自動的に更新するように設定されているか否かを判定し、自動的に更新するように設定されていないと判定した場合、ステップS93に進む。一方、ステップS92において、自動的に更新するように設定されていると判定された場合、ステップS93の処理はスキップされる。

【0194】

ステップS93において、エージェントプログラム11の提示部22は、データベースの更新が必要である旨をユーザに通知するとともに、さらに、その通知に対応して、ユーザから更新の指示がなされたか否かを判定する。ユーザから更新の指示がなされたと判定された場合、処理はステップS94に進む。反対に、ユーザから更新の指示がなされないと判定された場合、処理はステップS91に戻り、以降の処理が繰り返し実行される。

【0195】

ステップS94において、エージェントプログラム11の蓄積部21は、データベースを更新する。具体的には、文書取得部31乃至文書内容処理部33が、電子メールの電子メールボックスファイル（特定の拡張子mbx等が付与されていることが多い）を検出し、その更新日時を取得して、以前に取得した更新日時と比較し、異なる日付と異なるファイルサイズであれば、ファイルが更新されないと判断し、追加または変更された部分を抽出する。この場合、電子メールのグループ化、ヘッダの解析、形態素解析、特徴ベクトル算出等、一連のファイル内の分析が行われ、得られる重要語が関連情報検索部35に供給される。

【0196】

ただし、メールグループ（話題）が変化せず（所定の話題に新たに追加された電子メールがなく）、分析の結果、更新以前の重要語（検索用キーワード）と更新後の重要語が同じであれば、評価値等の計算値だけを変更し、関連情報検索部35による関連情報の検索を実行しないようにしてもよい。

【0197】

あるいは、全ての電子メールグループが変化せずに一定期間が経過した場合、グループの特徴ベクトルのうち、前回、評価値が1番目と2番目の単語を検索語としていたものを、例えば評価値が3番目と4番目の単語を検索語に変更して検索し、検索結果を取得するようにしてもよい。

【0198】

以上説明したように、データベース更新処理においては、追加された文書や変更された文書だけを処理の対象とするので、データベース作成処理を繰り返し実行する場合に比べて、処理時間が短縮される。

【0199】

次に、サーバ6のサーバプログラム101によるデータベース作成処理について、図23のフローチャートを参照して説明する。このデータベース作成処理は、サーバプログラム101が実行する処理のうちの1つであり、サーバプログラム101が起動された状態において、データベースが未だ作成されていないときに開始される。

【0200】

ステップS101において、番組メタデータ取得部121は、データベース作成の素として分析するEPGデータなどの番組情報（番組メタデータ）を取得し、データ内容処理部122に供給する。

【0201】

図24は、ステップS101において取得される番組メタデータ220の一例を示している。番組メタデータ220は、当該番組の番組名を示すタイトル221、当該番組の分類（例えば、ドラマ、映画、ニュース、スポーツ、あるいは音楽など）を示すジャンル222、当該番組の放送時間（例えば、朝、昼、夕方、ゴールデン、あるいは深夜など）を示す時間帯情報223、当該番組が放送されるチャンネル（例えば、NHK総合、日本テレビ、あるいはTBS（いずれも商標）など）を示す放送局224、当該番組に出演している人物を示す出演者情報225、当該番組の脚本、原作、あるいは演出した人物を示す脚本・原作・演出情報226、および、当該番組のあらすじや見所などを示す内容（キーワード）情報227から構成される。

【0202】

図23に戻る。ステップS102において、データ内容処理部122は、ステップS101の処理で取得された番組メタデータ220に形態素解析を施し、番組情報（番組名、ジャンル名、放送局名、時間帯情報、出演者名、およびキーワードなど）を抽出する。ステップS103において、データベース作成部123は、データ内容処理部122によって抽出された番組情報を番組毎にデータベース化して、記憶部147に記録する。

【0203】

以上のデータベース作成処理が実行されることにより、番組メタデータ220から抽出された番組情報が推薦番組用データベース内に蓄積されることになる。また、データベース作成処理をユーザが強制的に中断することができるようるために、中断要求があった場合、中断された時点で処理済みの文書を記録し、再開要求があった場合、未処理の文書から処理を再開するようにしてもよい。

【0204】

なお、データベース作成処理は、サーバプログラム101が実行された場合に開始されるものとしたが、任意のタイミングで開始させることができるとともに、所定の期間毎（例えば、EPGデータ更新毎）に更新される。

【0205】

次に、図7の処理でパーソナルコンピュータ1のデータベースに記録された興味データ211にマッチングする番組情報を、図23の処理で作成されたサーバ6の推薦用番組データベースから検索する処理について、図25のフローチャートを参照して説明する。

【0206】

ステップS121において、HDDレコーダ2のCPU71は、ユーザによって入力部76が操作され、ログインされたか否かを判定し、ログインされたと判定されるまで待機する。ステップS121において、ユーザによりログインされたと判定された場合、ステップS122に進み、CPU71は、興味データを取得するためのコマンドとともにユーザのログイン情報を、通信部77を介してパーソナルコンピュータ1に送信する。

【0207】

ステップS111において、パーソナルコンピュータ1のデータベース問い合わせ部42は、HDDレコーダ2から送信されてきた興味データ取得コマンドを受信し、文書特徴データベース作成部34によって作成されたデータベースを検索し、ログイン情報（ログインアカウントやパスワードなど）に対応する、興味データ211（図20）を読み出し、通信部60を介してHDDレコーダ2に送信する。

【0208】

ステップS123において、HDDレコーダ2のCPU71は、ログインされたユーザーに対応する興味データ211を、パーソナルコンピュータ1から受信し、RAM73に記録させる。ステップS124において、CPU71は、受信した興味データ211を、通信部77およびネットワーク5を介してサーバ6に送信し、興味データ211にマッチする番組情報の検索を要求する。

【0209】

ステップS131において、サーバ6のイベント管理部131は、ネットワーク5を介してHDDレコーダ2から送信されてきた興味データ211を受信する。そして、イベント管理部131は、興味データ211をデータベース問い合わせ部132に供給して検索要求を通知する。ステップS132において、データベース問い合わせ部132は、イベント管理部131からの検索要求の通知に対応して、データベース作成部123によって作成された推薦番組用データベースを検索し、検索要求に含まれる興味データ211にマッチする番組情報を推薦用番組として選択する。

【0210】

ここで、興味データ211を構成する興味語ベクトル212には、出演者のニックネームなどの未定義語、表記ゆれや短縮語、あるいは、言い回し（ニュアンス）の違いなどの単語が含まれている場合がある。従って、サーバ6では、これらの単語にマッチする番組情報を検索することができるようにするため、未定義語、表記ゆれ、および短縮語などの対応辞書を持つようになることが好ましい。

【0211】

ステップS133において、番組情報出力部133は、データベース問い合わせ部132で選択された推薦用番組（番組情報）を検索結果として、通信部148およびネットワーク5を介してHDDレコーダ2に送信する。

【0212】

ステップS125において、HDDレコーダ2のCPU71は、ネットワーク5を介してサーバ6から送信されてきた番組情報に含まれる録画開始時刻、録画終了時刻、およびチャンネルなどに基づいて、RAM73に番組の録画予約を設定するとともに、チューナ79、復調部80、デコーダ81、およびHDD78などを制御する。これにより、録画予約の開始時刻になると、RAM73に設定されている番組情報が読み出され、録画予約が実行される。

【0213】

以上のようにして、ユーザの興味にマッチした番組情報が検索され、HDDレコーダ2に録画予約が自動的に設定される。すなわち、ユーザがやりとりした電子メールから抽出した興味語を基に番組情報を検索するため、日常の興味が反映されたタイムリーな番組推薦が可能となる。

【0214】

またステップS111の処理でデータベースから興味データ211を読み出して送信する際に、興味データ211の興味語ベクトル212に含まれる各単語（興味語）のうち、評価値の時系列推移に着目して、所定条件を満たす単語（重要語）のみを選択するようにしてもよい。ここで所定の選択条件としては、例えば、条件1を「現時点以前の所定の期間Y（例えば、5週間）において、異なる2以上の話題で、当該単語の評価値が所定の閾値B以上であること」、条件2を「条件1における異なる2以上の話題のうち、最も古い話題と最も新しい話題が所定の期間Z以上離れていること」などとする。

【0215】

このような条件を用いることにより、ユーザが高い関心を持っていると思われる単語（重要語）や、ユーザが現時点で意識していない意外性があると思われる単語にマッチした番組情報を推薦することが可能となる。

【0216】

なお、上述した例においては、電子メールの形態素解析、およびその解析結果から抽出される興味データ211の作成処理をパーソナルコンピュータ1で行い、HDDレコーダ2には、興味データ211のみを送信させるようにしたが、HDDレコーダ2に、パーソナルコンピュータ1のエージェントプログラム11と同等の機能を有するプログラムを実装させようすれば、電子メールの形態素解析や興味データ211の作成処理をHDDレコーダ2に行わせることも可能である。

【0217】

また、パーソナルコンピュータ1では、電子メールの形態素解析、および解析結果に基づいて興味データ211を作成し、HDDレコーダ2では、パーソナルコンピュータ1で作成された興味データ211を取得してサーバ6に送信し、興味データ211にマッチする番組情報の推薦を受けるようにし、サーバ6では、興味データ211にマッチする番組情報を検索し、検索結果をHDDレコーダ2に送信するようにして、処理の分散を図るようにした。

【0218】

本発明は、これに限らず、例えば、パーソナルコンピュータ1が、サーバ6で作成された推薦番組用データベースの番組情報を一括して提供してもらい、番組情報の検索を行うこともできる。その場合、HDDレコーダ2は、パーソナルコンピュータ1から供給される番組情報に含まれる録画開始時刻、録画終了時刻、およびチャンネルなどを取得し、録画予約を実行する。

【0219】

さらに、サーバ6が、例えば、1つのインターネットサービスプロバイダであり、上述した推薦番組の配信サービスやメールサービスを行っている場合、ユーザのメールサーバに直接アクセスして、サーバ6側で電子メールの形態素解析、およびその解析結果に基づいて興味データ211を作成するようにしてもよい。

【0220】

ところで、上述した処理では、ユーザに気付かれずに、ユーザの興味にマッチした番組が検索され、自動的にその番組の録画予約が実行されてしまう。そこで、例えば、HDDレコーダ2の電源オン時、録画予約確認時、あるいは、録画番組の再生時に、録画予約された番組の推薦理由を提示させるようにする。その場合

の処理について、図26のフローチャートを参照して説明する。

【0221】

ステップS141において、HDDレコーダ2のCPU71は、提示条件を満たすか否かを判定し、提示条件を満たすと判定されるまで待機する。ここで提示条件とは、例えば、HDDレコーダ2に番組の録画予約が設定され、かつ、テレビジョン受像機3が電源オン状態であるか、あるいは、HDDレコーダ2に番組の録画予約が設定され、かつ、HDDレコーダ2のGUI画面が表示可能状態であるなどである。

【0222】

ステップS141において、提示条件を満たすと判定された場合、ステップS142に進み、CPU71は、図25のステップS123の処理でRAM73に記録された興味データ211に含まれるトピックスID162乃至メンバー情報165(図20)に基づいて、推薦理由(背景情報)をテレビジョン受像機3に出力し、表示させる。

【0223】

図27は、推薦理由の表示例を示している。同図に示されるように、エージェント231を登場させるとともに、エージェント231の台詞として吹き出し232、および、ユーザにより操作可能な入力ウィンドウ233が表示される。吹き出し232の中には、例えば「2001年4月9日、太郎さんと「ワインありがとう」ってメールしたよね。このまま録画する?」の台詞が表示される。

【0224】

また、吹き出し232の表示と同期して、吹き出し232に表示された台詞と同じ意味を持つ言語(例えば、日本語や英語など)の音声信号が音声合成部(図示せず)によって合成されて出力させるようにすることができる。ただし、吹き出し232の表示の有無や台詞に対応する音声の出力の有無はエージェントプログラム11が適宜設定するか、ユーザが任意に設定できるようにすることが可能である。

【0225】

また入力ウィンドウ233には、番組を録画予約するときに選択する「録画す

る」ボタン、番組の録画予約をキャンセルするときに選択する「キャンセル」ボタンが表示される。

【0226】

ユーザは、吹き出し232に表示された台詞によって、「太郎さん」とメールした内容（いまの場合、「ワインありがとう」）から、「△△△のワイン紀行」の番組がサーバ6から推薦され、HDDレコーダ2に録画予約されたことを知ることができる。そして、ユーザは、このまま録画を開始させる場合には、「録画する」ボタンを選択し、録画をキャンセルさせる場合には、「キャンセル」ボタンを選択する。なお、ユーザによりいずれのボタンも選択されなかった場合を想定して、タイムアウトをトリガにして、いずれかのボタンが自動的に選択されるようにもよい。

【0227】

図26に戻る。ステップS143において、CPU71は、ユーザによって入力部76が操作されて「キャンセル」ボタンが選択されたか否かを判定し、「キャンセル」ボタンが選択されたと判定した場合、ステップS144に進み、RAM73に設定されている録画予約を削除させる。

【0228】

ステップS143において、「キャンセル」ボタンが選択されなかった、すなわち、「録画する」ボタンが選択されたと判定された場合、ステップS144の処理はスキップされる。

【0229】

以上のようにして、自動的に録画予約された番組の推薦理由を表示させることができ、その番組を録画するか否かをユーザに決定させることができる。

【0230】

また上述した例では、HDDレコーダ2からテレビジョン受像機3に推薦理由を表示させるようにしたが、これに限らず、例えば、図25のステップS133の処理において、サーバ6がHDDレコーダ2に検索結果を送信する際に、パーソナルコンピュータ1にも検索結果を送信するようにし、パーソナルコンピュータ1の電源オン時、あるいは、エージェントプログラム11の起動時に、推薦理由を

表示させることもできる。その場合の推薦理由の表示例を図28に示す。

【0231】

図28の例の場合、吹き出し232の中には、例えば「2001年4月9日、太郎さんとメールした「ワインありがとう」から「△△△のワイン紀行」という番組の推薦があったよ。録画する？」の台詞が表示される。

【0232】

ユーザは、吹き出し232に表示された台詞によって、「太郎さん」とメールした内容（いまの場合、「ワインありがとう」）から、「△△△のワイン紀行」の番組がサーバ6から推薦され、HDDレコーダ2に録画予約されたことを知ることができる。そして、ユーザは、このまま録画を開始させる場合には、「録画する」ボタンを選択し、録画をキャンセルさせる場合には、「キャンセル」ボタンを選択する。「キャンセル」ボタンが選択された場合、パソコンコンピュータ1のCPU41は、HDDレコーダ2に、録画予約の削除コマンドを送信し、録画予約の削除を実行させる。

【0233】

以上のようにして、自動的に録画予約された番組の推薦理由を、HDDレコーダ2を介してテレビジョン受像機3に、またはパソコンコンピュータ1に表示させることができる。また推薦理由には、ユーザが電子メールでやりとりした話題、日時、やりとりした相手、またはサブジェクトなどが含まれているため、単に「××番組が推薦されました」という提示に較べ、より興味を引くことができる。

【0234】

以上においては、ユーザがやりとりした電子メールや作成した文書などに含まれる単語（興味語）にマッチした番組情報を検索し、推薦するようにしたが、これに限らず、例えば、ユーザが初期登録した嗜好情報（好みのジャンル名や出演者名など）、または、ユーザが視聴した番組の視聴履歴などを分析し、その分析結果から興味語を抽出して、マッチングする番組情報を検索することもできる。その場合の処理について、図29のフローチャートを参照して説明する。

【0235】

ステップS151において、パーソナルコンピュータ1のイベント管理部41は、内蔵するタイマ41Aを参照することによって時間の経過を監視し、所定期間が経過したか否かを判定し、所定期間が経過するまで待機する。そして、所定期間が経過したと判定された場合、イベント管理部41は、イベント発生をデータベース問い合わせ部42に通知し、ステップS152に進む。

【0236】

ステップS152において、データベース問い合わせ部42は、文書特徴データベース作成部34によって作成されたデータベースから、初期登録された嗜好情報または視聴履歴が形態素解析されて記録された興味データ211を読み出し、ネットワーク5を介してサーバ6に送信する。

【0237】

ステップS171において、サーバ6のイベント管理部131は、ネットワーク5を介してパーソナルコンピュータ1から興味データ211を受信する。そして、イベント管理部131は、興味データ211をデータベース問い合わせ部132に供給して検索要求を通知する。ステップS172において、データベース問い合わせ部132は、イベント管理部131からの検索要求の通知に対応して、データベース作成部123によって作成された推薦番組用データベースを検索し、検索要求に含まれる興味データにマッチする番組情報を推薦用番組として選択する。

【0238】

ステップS173において、番組情報出力部133は、データベース問い合わせ部132で選択された推薦用番組（番組情報）を検索結果として、通信部148およびネットワーク5を介してパーソナルコンピュータ1およびHDDレコーダ2に送信する。

【0239】

ステップS161において、HDDレコーダ2のCPU71は、ネットワーク5を介してサーバ6から送信されてきた番組情報に含まれる録画開始時刻、録画終了時刻、およびチャンネルなどに基づいて、RAM73に番組の録画予約を設定するとともに、チューナ79、復調部80、デコーダ81、およびHDD78などを制御

する。これにより、録画予約の開始時刻になると、RAM73に設定されている番組情報が読み出されて、録画予約が実行される。

【0240】

ステップS153において、イベント管理部41は、検索結果を受信し、イベント発生をデータベース問い合わせ部32に通知する。データベース問い合わせ部42は、イベント管理部41からのイベント発生の通知に対応して、イベント発生に対応する検索結果（番組情報）を取得し、形態素解析を施して単語（特徴語）を抽出し、各単語の評価値を演算する。これにより、検索結果（番組情報）の特徴ベクトルが算出される。

【0241】

ステップS154において、データベース問い合わせ部42は、文書特徴データベース作成部34が作成したデータベースを検索し、ステップS153の処理で算出された特徴ベクトルと、データベースに記録されている話題毎の特徴ベクトルとの内積を両者の類似度として算出し、類似度が所定の条件（例えば、類似度が最大、もしくは類似度が所定の閾値以上）を満たす話題を抽出する。

【0242】

このとき、例えば、特定のジャンル名（ユーザによって初期登録された番組のジャンル名）を含む話題を予め選抜しておくようにして、検索結果（番組情報）に類似する話題を効率良く抽出することができるようにしてよい。また、一般的な番組のジャンル名を含む話題を一通り選抜しておくようにしてもよい。

【0243】

ステップS155において、データベース問い合わせ部42は、ステップS154の処理で抽出された話題を構成する文書のうち、最新の文書を選択し、イベント管理部41を介して、または直接的に、関連情報提示部43に供給する。ステップS156において、エージェント制御部23は、ステップS155の処理で選択した文書の属性情報を、選択（推薦）した理由としてデスクトップに表示させる（図28）。

【0244】

ステップS157において、エージェントプログラム11は、入力部56が操

作されて「キャンセル」ボタンが選択されたか否かを判定し、「キャンセル」ボタンが選択されたと判定された場合、ステップS158に進み、録画予約の削除コマンドをHDDレコーダ2に送信する。ステップS157において、「キャンセル」ボタンが選択されなかった、すなわち、「録画する」ボタンが選択されたと判定された場合、ステップS158の処理はスキップされる。

【0245】

ステップS162において、HDDレコーダ2のCPU71は、パーソナルコンピュータ1から送信されてきた録画予約の削除コマンドを受信し、ステップS163において、RAM73に設定されている録画予約を削除させる。

【0246】

以上のようにして、ユーザによって初期登録された嗜好情報や視聴履歴などの多くの情報を形態素解析して、有意な興味データ211を作成するため、その興味データ211にマッチした、意外性のある番組情報を推薦することができる。また、推薦番組に類似する話題（または文書）を文書特徴データベース作成部34が作成したデータベースから逆引きすることで、推薦された番組に関連する推薦理由を取得することができる。従って、ユーザは、個人的な結びつきによって番組が推薦されたかのように感じることができる。

【0247】

以上においては、ユーザがやりとりした電子メールや作成した文書、あるいは嗜好情報や視聴履歴から興味データ211を作成し、それにマッチした番組情報を検索し、推薦するようにした。すなわち、パーソナルコンピュータ1またはHDDレコーダ2は、予め作成された興味データ211をデータベースから読み出し、ネットワーク5を介してサーバ6に送信することによって、マッチングする番組の推薦を受けるようにした。従って、ユーザの意思に関わらず、ユーザの興味があると思われる番組を自動的に検索し、推薦するというものである。

【0248】

ところで、当然のことながら、ユーザの意思によって、番組の録画予約を行うようにすることもできる。次に、図30のフローチャートを参照して、ユーザが、例えば、外出先において、所有しているデジタル携帯電話機4を用いて、番組

の録画予約を行う場合の処理について説明する。

【0249】

ユーザは、デジタル携帯電話機4を操作し、図31に示されるような、番組の録画予約をするための電子メール（以下、予約メールと記述する）241を作成する。同図に示されるように、予約メール241の本文には、「今日、音楽番組を録画して。特にクラシックものは絶対。あればジャズ番組も。あとワールドカップ情報もお願い。」といったメッセージが記載されている。

【0250】

すなわち、ユーザは、予約メール241を作成するにあたって、記述フォーマットを気にする必要なく、自由文で作成すればよく、予約メール241には、録画予約したい番組の番組名や出演者名などの一部、あるいは、ジャンル名を入力するだけでもよい。ここで、予約メール241の受信側において、その電子メールが番組を録画予約するための電子メール（すなわち、予約メール241）であることを認識することができるよう、例えば、サブジェクト（表題）に、「予約メール」あるいは「録画予約」といった定型文（図31の例の場合、「予約メール」）を入力するようにする。

【0251】

ユーザは、作成した予約メール（図31）を、自宅にあるパーソナルコンピュータ1に送信するように、デジタル携帯電話機4を操作する。これにより、ステップS1.81において、デジタル携帯電話機4は、ユーザの操作に対応する入力信号を受けて、作成された予約メール241を、ネットワーク5を介してパーソナルコンピュータ1に送信する。

【0252】

ステップS1.91において、パーソナルコンピュータ1のイベント管理部41は、ネットワーク5を介して予約メール241を受信したことを検知して、データベース問い合わせ部42に通知する。ステップS1.92において、データベース問い合わせ部42は、イベント管理部41からのイベント発生に対応する予約メール241を取得する。このときデータベース問い合わせ部42は、予約メール241のサブジェクトに「予約メール」と記述されていることにより、番組を

録画予約するための電子メールであると認識する。そしてデータベース問い合わせ部42は、取得した予約メール241に形態素解析を施して単語を抽出し、不要語を除外して興味語ベクトル（特徴ベクトル）を作成（算出）する。

【0253】

いまの場合、予約メール241に記載された「今日、音楽番組を録画して。特にクラシックものは絶対。あればジャズ番組も。あとワールドカップ情報もお願い。」から、形態素解析によって、「今日、音楽、クラシック、ジャズ、ワールドカップ」が興味語ベクトルとして抽出される。また、「今日」はこの時点での時情報（例えば、2001年4月9日）に変換される。

【0254】

なお、カテゴリに木構造を予め定義しておくことによって、形態素解析の分析精度を向上させることができる。例えば、音楽のカテゴリに{クラシック、ジャズ、ポップス、ロック、演歌、・・・}などと定義しておき、形態素解析によって抽出された興味語ベクトルに適用させると、「クラシック、ジャズ、ワールドカップ」の3つの単語（興味語）で興味語ベクトルが構成される。

【0255】

ステップS193において、データベース問い合わせ部42は、ステップS192の処理で作成された興味語ベクトルを、ネットワーク5を介してサーバ6に送信する。

【0256】

ステップS211において、サーバ6のイベント管理部131は、ネットワーク5を介してパーソナルコンピュータ1から興味語ベクトルを受信する。そして、イベント管理部131は、興味語ベクトルをデータベース問い合わせ部132に供給して検索要求を通知する。ステップS212において、データベース問い合わせ部132は、イベント管理部131からの検索要求の通知に対応して、データベース作成部123によって作成された推薦番組用データベースを検索し、検索要求に含まれる興味語ベクトルにマッチする番組情報を録画予約用の番組として選択する。

【0257】

ステップS213において、番組情報出力部133は、データベース問い合わせ部132で選択された録画予約用の番組（番組情報）を検索結果として、ネットワーク5を介してHDDレコーダ2に送信する。

【0258】

ステップS201において、HDDレコーダ2のCPU71は、ネットワーク5を介してサーバ6から送信されてきた番組情報に含まれる録画開始時刻、録画終了時刻、およびチャンネルなどに基づいて、RAM73に番組の録画予約を設定するとともに、チューナ79、復調部80、デコーダ81、およびHDD78などを制御する。これにより、録画予約の開始時刻になると、RAM73に設定されている番組情報が読み出され、録画予約が実行される。

【0259】

ステップS202において、CPU71は、録画予約の設定が完了すると、例えば、図32に示されるような、番組の録画予約を完了した旨を通知するための電子メール（以下、予約完了メールと記述する）251を作成する。同図に示されるように、予約完了メール251の本文には、「下記の番組を録画しました」といったメッセージとともに、第1番目のエントリに、「4CH」の放送局の「19:00～20:00」の間に「ワールドカップハイライト」が録画予約された旨が記載され、第2番目のエントリに、「3CH」の放送局の「21:00～21:54」の間に「×××クラシック」が録画予約された旨が記載される。

【0260】

そして、作成された予約完了メール251は、ネットワーク5を介してデジタル携帯電話機4に送信される。

【0261】

ステップS182において、デジタル携帯電話機4は、ネットワーク5を介してHDDレコーダ2より送信されてきた予約完了メール251を受信する。そして、ユーザが電子メールを表示させるように操作すると、ステップS183において、デジタル携帯電話機4は、ユーザの操作に対応する入力信号に基づいて、受信した予約完了メール251をディスプレイに表示させる。

【0262】

以上のようにして、パーソナルコンピュータ1のエージェントプログラム11は、送信されてきた予約メール241を形態素解析し、興味語ベクトル（録画予約に必要な情報）を作成してサーバ6に送信する。サーバ6は、パーソナルコンピュータ1から送信されてきた興味語ベクトルにマッチする所定の番組情報を検索し、HDDレコーダ2に送信する。これにより、HDDレコーダ2には、自動的に番組の録画予約が行われる。

【0263】

従って、ユーザは、外出先においても、記述形式を気にすることなく自由文で作成した予約メール241を自宅のパーソナルコンピュータ1に送信するだけでも、簡単に番組の録画予約を行うことができるとともに、番組の録画予約が完了したことを簡単に知ることもできる。

【0264】

以上説明した本実施の形態においては、ユーザがやりとりした電子メールの内容、ユーザによって初期登録された嗜好情報、および視聴履歴などを形態素解析して興味データ211を作成し、その興味データ211にマッチする番組情報の推薦を受けるようにするか、あるいは、ユーザが作成した予約メール241（番組を録画予約するための電子メール）を形態素解析して興味語ベクトルを作成し、その興味語ベクトルにマッチする番組情報を録画するようにした。従って、ユーザは、日常的な関心からの番組や潜在的な関心からの番組といったように、場面や目的に応じた多様な番組推薦を受けることができる。

【0265】

また、サーバ6は、強調フィルタリング方式を用いて、他のユーザの興味データ211を利用して番組推薦を行うこともできる。その場合、パーソナルコンピュータ1またはHDDレコーダ2は、興味データ211に含まれる情報のうち、プライバシー上問題の無いデータになるようにフィルタリングした後（例えば、番組名は選択せずに、ジャンル名のみを選択した後）、サーバ6に送信するようとする。このフィルタリングのレベルは、ユーザが任意に設定することができる。

【0266】

さらに、図27の表示例を用いて説明したように、推薦理由の提示時、推薦さ

れた番組を録画予約するか否かをユーザにより決定させるため、HDDレコーダ2は、録画予約が実行された番組と、録画予約がキャンセルされた番組に分類して、推薦番組のアクセス履歴を保存することができる。従って、HDDレコーダ2は、推薦番組を録画予約する頻度をカウントしたり、あるいは、アクセス履歴に形態素解析を施して新たな興味を抽出することにより、例えば、図33に示されるような嗜好データ260を作成することが可能になる。

【0267】

図33の例の場合、嗜好データ260は、推薦番組の分類（例えば、ドラマ、映画、ニュース、スポーツ、あるいは音楽など）を示すジャンル261、当該推薦番組の番組名を示すタイトル262、当該推薦番組の放送時間（例えば、朝、昼、夕方、ゴールデン、あるいは深夜など）を示す時間帯情報263、当該推薦番組に出演している人物を示す出演者情報264、当該推薦番組のあらすじや見所などを示す内容（キーワード）情報265、および、当該推薦番組と一緒に視聴していたユーザに関する同席者情報266から構成される。

【0268】

なお、同席者情報266の作成方法は、例えば、ユーザが入力部57を用いて入力するようにしてもよいし、CPU41が、各ユーザが所有している、デジタル携帯電話機や時計などから常時発信される機器IDを検知し、機器IDからユーザを特定して入力するようにしてもよいし、あるいは、各ユーザの会話から音声認識し、認識結果からユーザを特定して入力するようにしてもよい。

【0269】

このようにして作成された嗜好データ260は、ユーザが任意にプライバシー上問題の無いデータになるようにフィルタリングされてサーバ6に送信される。そして、嗜好データ260を受信したサーバ6は、その嗜好データ260にマッチする番組情報を検索し、ユーザが新たに興味を引くような番組を推薦することができる。

【0270】

なお、図20に示した興味データ211、および図33に示した嗜好データ260は、所定のスキーマを定義し、例えば、拡張性のあるXML (eXtensible Mark

up Language) などで記述することによって、HTTP (Hypertext Transport Protocol) を用いて、任意の場所および任意の機器から、容易に、アクセスすることができるようにもよい。

【0271】

また、パーソナルコンピュータ1とHDDレコーダ2は、Ethernet(R)を介してデータの授受が行われるものとして説明したが、例えば、i.Link(登録商標)、またはIEEE(Institute of Electrical and Electronics Engineers)802.11a、IEEE802.11b、Bluetooth(登録商標)などの無線LANを介してデータの授受を行うようにしてもよい。さらに、磁気ディスク、光ディスク、光磁気ディスク、もしくは半導体メモリよりなるリムーバブルメディアを用いてオフラインでデータを移動させてもよい。

【0272】

また、本実施の形態においては、電子メールから抽出された興味データにマッチする番組を推薦するようにしたが、これに限らず、例えば、ラジオ番組、またはインターネット上のWebサイト情報などを推薦することも勿論可能である。

【0273】

さらに本発明によれば、番組を推薦する際に、エージェント231によって推薦理由を提示させるようにしたので、エージェント231に対して信頼性や親しみが感じられるようになる。

【0274】

ところで、上述したエージェント231の表示および吹き出し232の中の台詞の表示、並びに、表示された台詞に対応する音声信号の出力については、本発明のエージェントプログラム11だけでなく、他のアプリケーション、例えば、ゲームやワードプロセッサのヘルプ画面等に適用することが可能である。さらに、ビデオカメラ、またはカーナビゲーション等のディスプレイに表示されるキャラクタに適用することも勿論可能である。

【0275】

上述した一連の処理を実行するエージェントプログラム11またはサーバプログラム101は、パーソナルコンピュータに予め組み込まれるか、あるいは、記

録媒体からインストールされる。

【0276】

上述した一連の処理は、ハードウェアに実行させることもできるが、通常、ソフトウェアにより実行させる。一連の処理をソフトウェアにより実行させる場合には、そのソフトウェアを構成するエージェントプログラム1が、専用のハードウェアに組み込まれているコンピュータ、または、各種のプログラムをインストールすることで、各種の機能を実行することが可能な、例えば汎用のパーソナルコンピュータなどに、記録媒体からインストールされる。

【0277】

コンピュータにインストールされ、コンピュータによって実行可能な状態とされるプログラムを記録する記録媒体は、図3または図6に示されるように、プログラムが記録されている磁気ディスク62, 150（フレキシブルディスクを含む）、光ディスク63, 151（CD-ROM（Compact Disk-Read Only Memory）, DVD（Digital Versatile Disk）を含む）、光磁気ディスク64, 152（MD（Mini-Disk）を含む）、もしくは半導体メモリ65, 153などよりなるパッケージメディア、または、プログラムが一時的もしくが永続的に記録されるROM52, 142や記憶部59, 147を構成するハードディスクなどにより構成される。記録媒体に対するプログラムの記録は、必要に応じてルータ、モデムなどのインターフェースを介して、公衆回線網、ローカルエリアネットワーク、インターネット、デジタル衛星放送といった、有線または無線の通信媒体を利用して行われる。

【0278】

なお、本明細書において、記録媒体に記録されるプログラムを記述するステップは、記載された順序に沿って時系列的に行われる処理はもちろん、必ずしも時系列的に処理されなくとも、並列的あるいは個別に実行される処理をも含むものである。

【0279】

また、本明細書において、システムとは、複数の装置により構成される装置全体を表わすものである。

【0280】

【発明の効果】

以上のように、第1の本発明によれば、テレビ番組の推薦を行うことができる

【0281】

また第1の本発明によれば、ユーザがやりとりした電子メールから興味を抽出し、その興味にマッチしたテレビ番組の推薦を行うことができる。

【0282】

第2の本発明によれば、容易にテレビ番組の検索を要求することができる。

【0283】

また第2の本発明によれば、ユーザがやりとりした電子メールから興味を抽出し、その興味にマッチしたテレビ番組の検索を要求することができる。

【0284】

第3の本発明によれば、データベース化された番組情報から、テレビ番組を検索することができる。

【0285】

また第3の本発明によれば、ユーザがやりとりした電子メールから抽出された興味に基づいて、その興味にマッチしたテレビ番組をデータベースから検索して推薦することができる。

【0286】

第4の本発明によれば、外出先においても簡単にテレビ番組の録画予約を行うことができる。

【0287】

また第4の本発明によれば、外出先においてテレビ番組の録画予約を行うための電子メールを作成し、自宅にある録画装置に送信する。そして、録画装置は、テレビ番組の録画予約を行うための電子メールを受信し、その電子メールから予約情報を抽出し、予約情報にマッチしたテレビ番組の番組情報をサーバから取得して録画予約を行うことができる。

【0288】

第5の本発明によれば、簡単にテレビ番組の録画予約を行うことができる。

【0289】

また第5の本発明によれば、テレビ番組の録画予約を行うための電子メールを受信し、その電子メールから予約情報を抽出し、予約情報にマッチしたテレビ番組の番組情報をサーバから取得して録画予約を行うことができる。

【図面の簡単な説明】

【図1】

本発明を適用した番組検索システムの構成例を示す図である。

【図2】

図1のパーソナルコンピュータのエージェントプログラムの機能を説明する図である。

【図3】

エージェントプログラムをインストールして実行させるパーソナルコンピュータの構成例を示すブロック図である。

【図4】

HDDレコーダの構成例を示すブロック図である。

【図5】

図1のサーバのサーバプログラムの機能を説明する図である。

【図6】

サーバプログラムをインストールして実行させるサーバの構成例を示すブロック図である。

【図7】

エージェントプログラムによるデータベース作成処理を説明するフローチャートである。

【図8】

図7のステップS1の処理を説明するフローチャートである。

【図9】

図8のステップS22の処理を説明するフローチャートである。

【図10】

話題ファイルの一例を示す図である。

【図1 1】

単語ベクトルを構成する複数の単語に含まれる要素を示す図である。

【図1 2】

図7のステップS 3の処理を説明するフローチャートである。

【図1 3】

図7のステップS 4の処理を説明するフローチャートである。

【図1 4】

話題単語テーブルの構成例を示す図である。

【図1 5】

単語インデックステーブルの構成例を示す図である。

【図1 6】

話題評価値テーブルの構成例を示す図である。

【図1 7】

図7のステップS 5の処理を説明するフローチャートである。

【図1 8】

図7のステップS 9の処理を説明するフローチャートである。

【図1 9】

図7のステップS 10の処理を説明するフローチャートである。

【図2 0】

興味データの一例を示す図である。

【図2 1】

データベース更新処理を説明するフローチャートである。

【図2 2】

データベースを更新させる条件を入力するユーザインターフェースの表示例を示す図である。

【図2 3】

サーバプログラムによるデータベース作成処理を説明するフローチャートである。

【図24】

番組メタデータの一例を示す図である。

【図25】

番組情報の検索処理を説明するフローチャートである。

【図26】

番組の推薦理由提示処理を説明するフローチャートである。

【図27】

推薦理由の表示例を示す図である。

【図28】

他の推薦理由の表示例を示す図である。

【図29】

番組情報の検索処理を説明するフローチャートである。

【図30】

番組の録画予約処理を説明するフローチャートである。

【図31】

予約メールの一例を示す図である。

【図32】

予約完了メールの一例を示す図である。

【図33】

嗜好データの一例を示す図である。

【符号の説明】

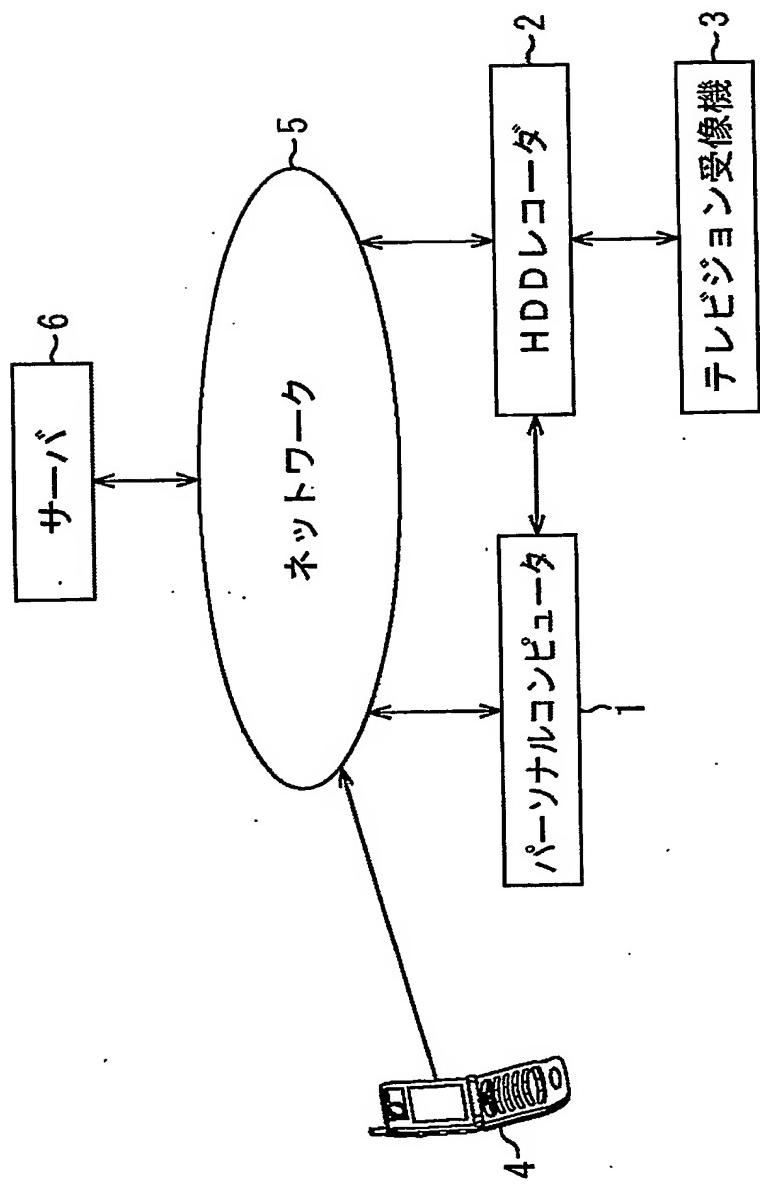
- 1 パーソナルコンピュータ, 2 HDDレコーダ, 3 テレビジョン受像機, 4 デジタル携帯電話機, 5 ネットワーク, 6 サーバ, 11 エージェントプログラム, 12 メーラ, 21 蓄積部, 22 提示部, 23 エージェント制御部, 31 文書取得部, 32 文書属性処理部, 33 文書内容処理部, 34 文書特徴データベース作成部, 35 関連情報検索部, 41 イベント管理部, 42 データベース問い合わせ部, 43 関連情報提示部, 101 サーバプログラム, 111 蓄積部, 1 12 検索部, 121 番組メタデータ取得部, 122 データ内容処理部

, 123 データベース作成部, 131 イベント管理部, 132 デー
タベース問い合わせ部, 133 番組情報出力部

【書類名】図面

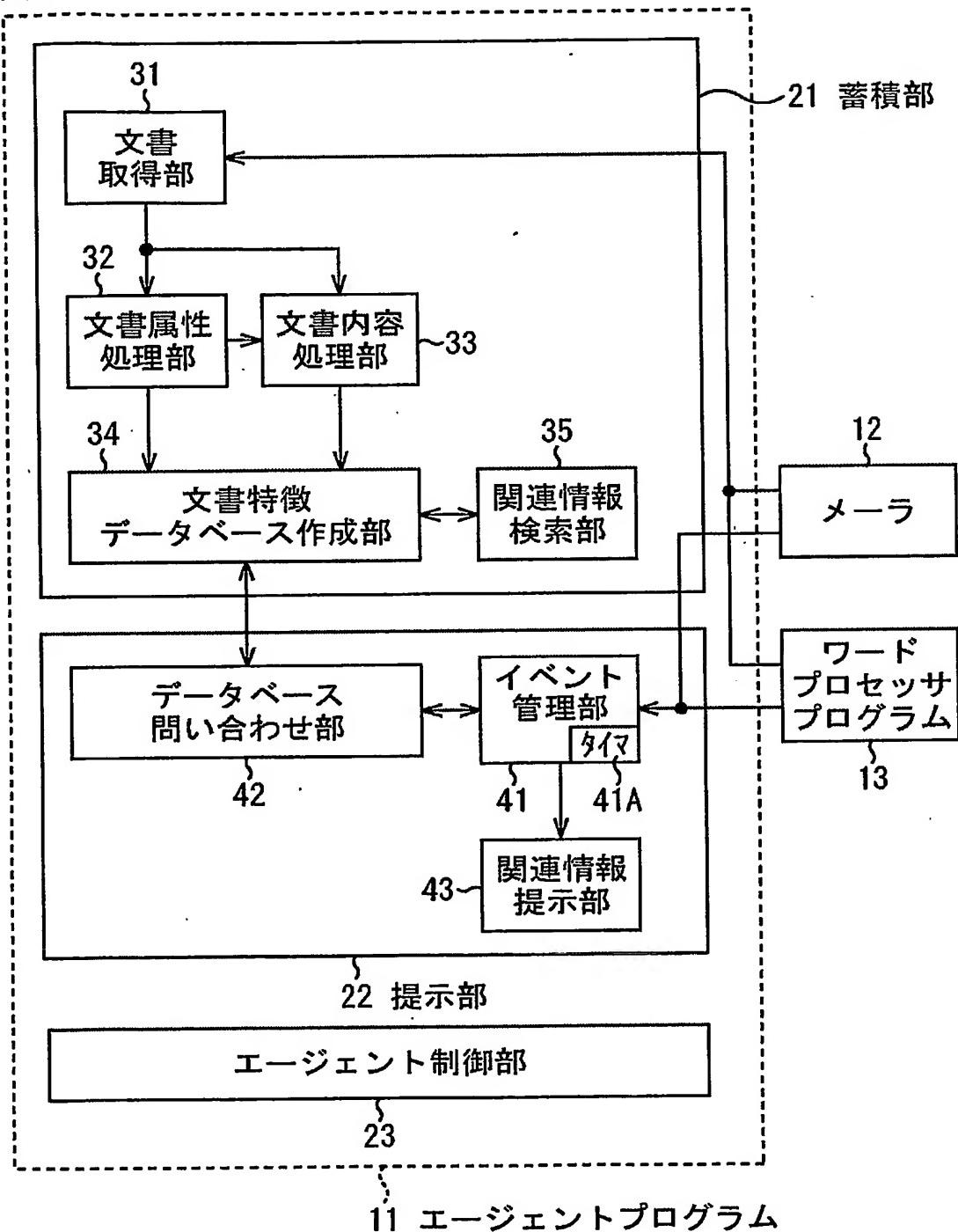
【図1】

図1



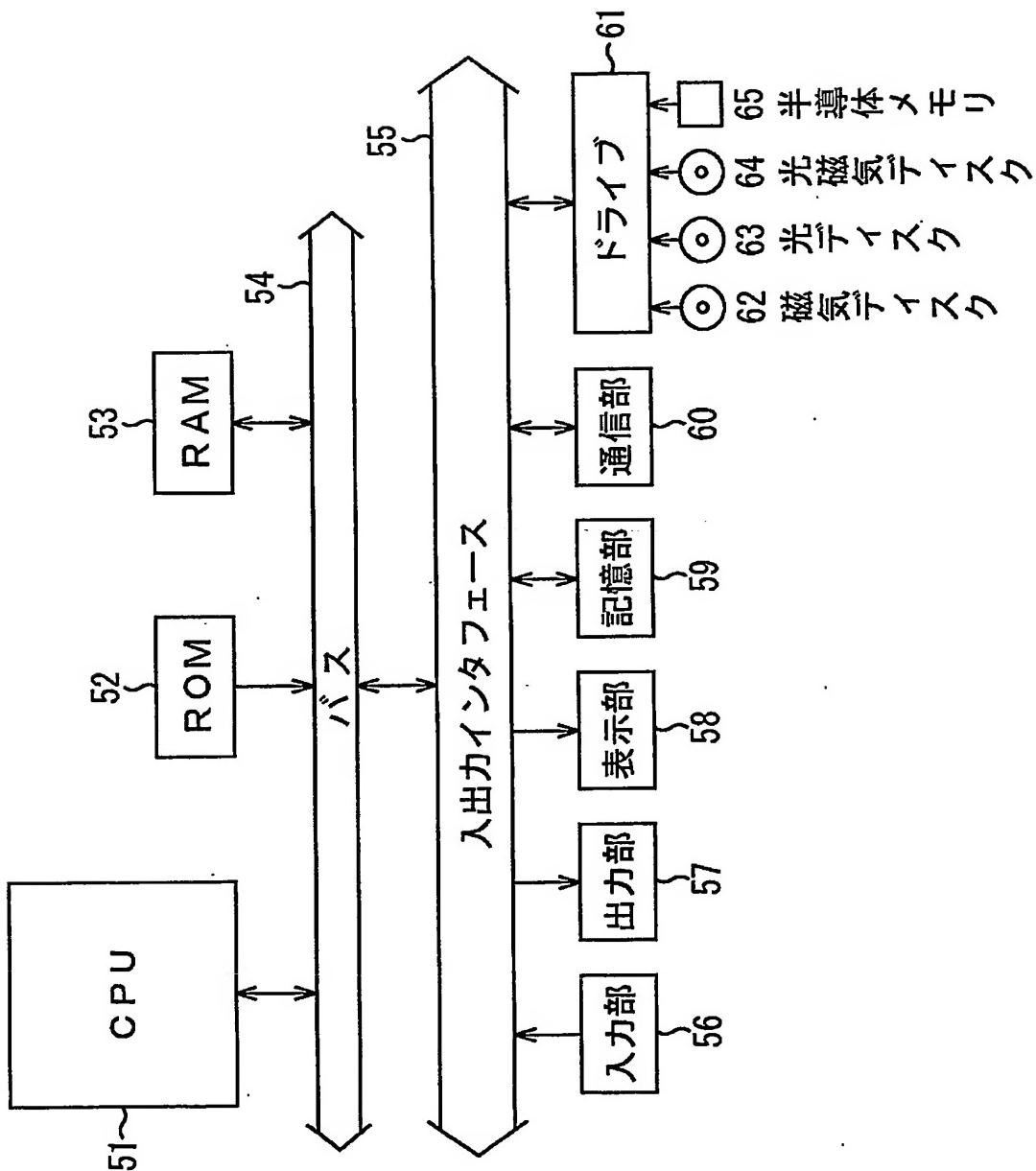
【図2】

図2



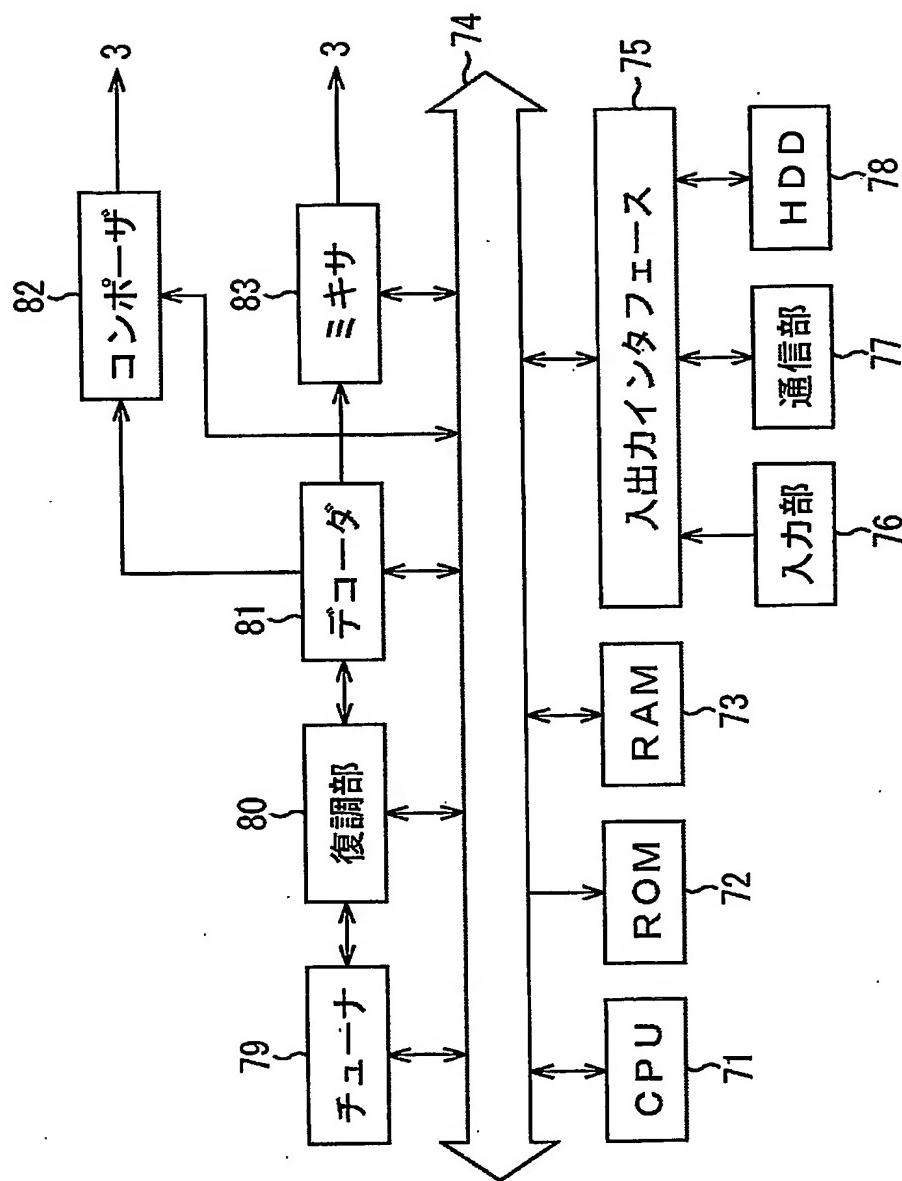
【図3】

図3



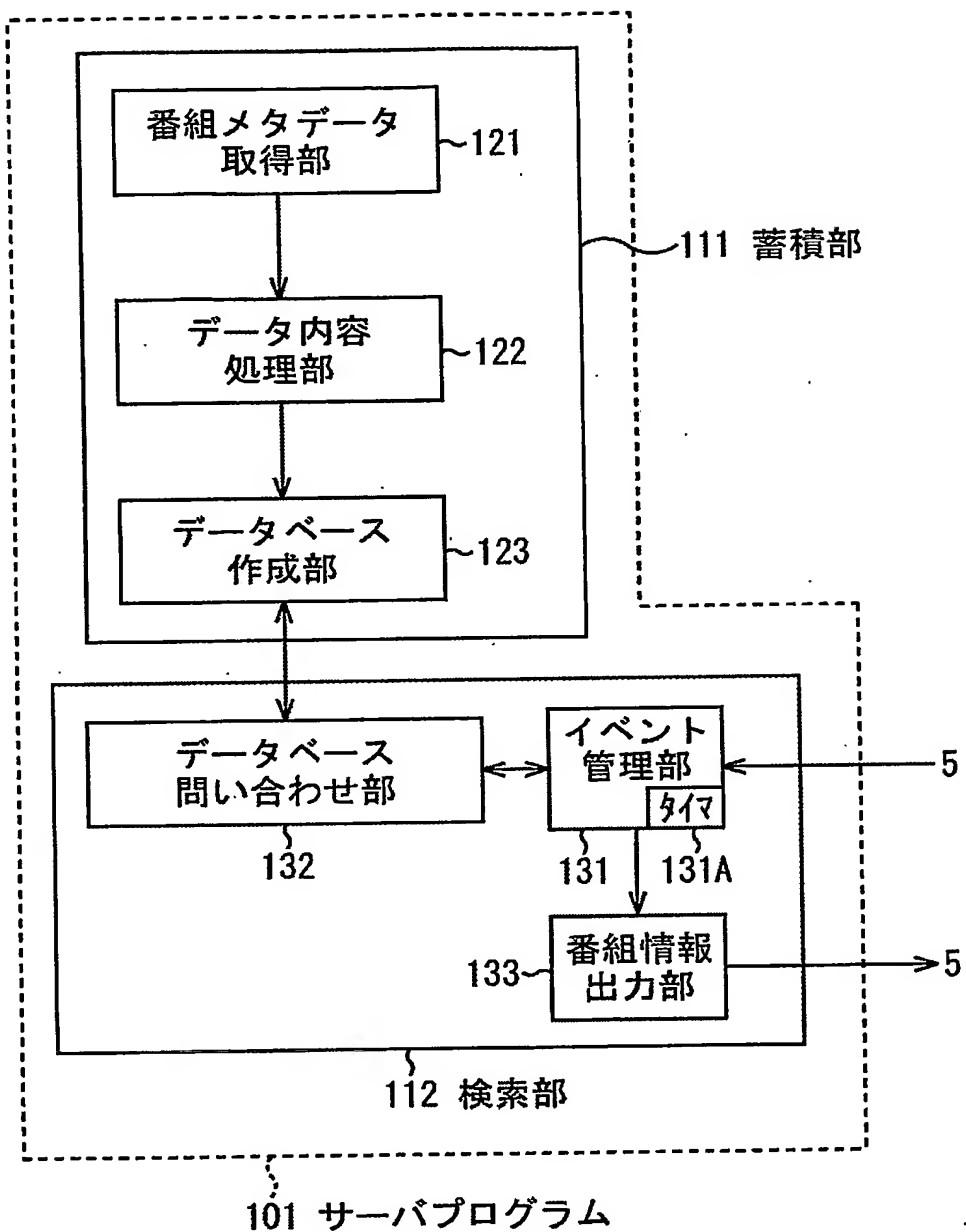
【図4】

図4



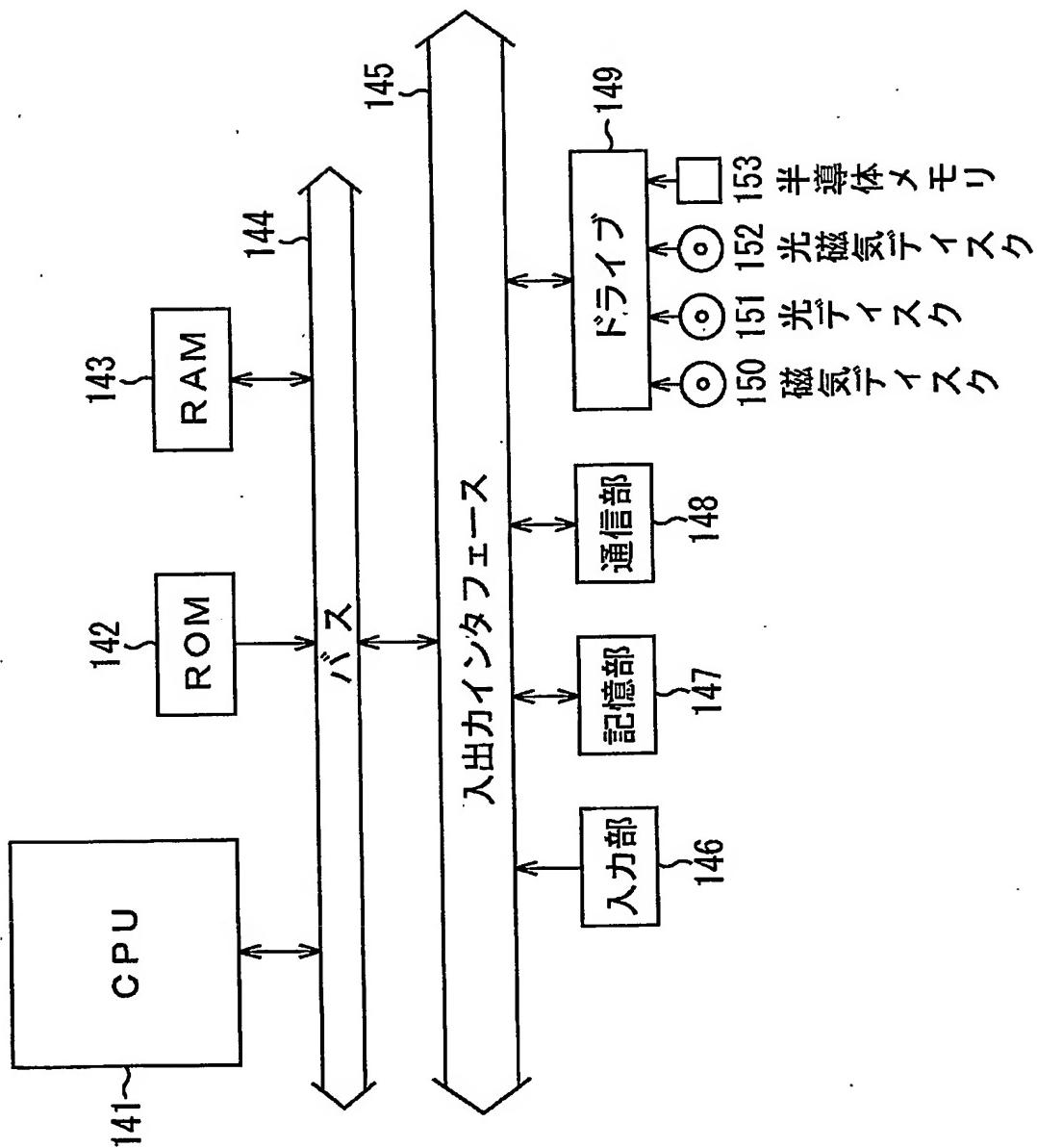
【図5】

図5



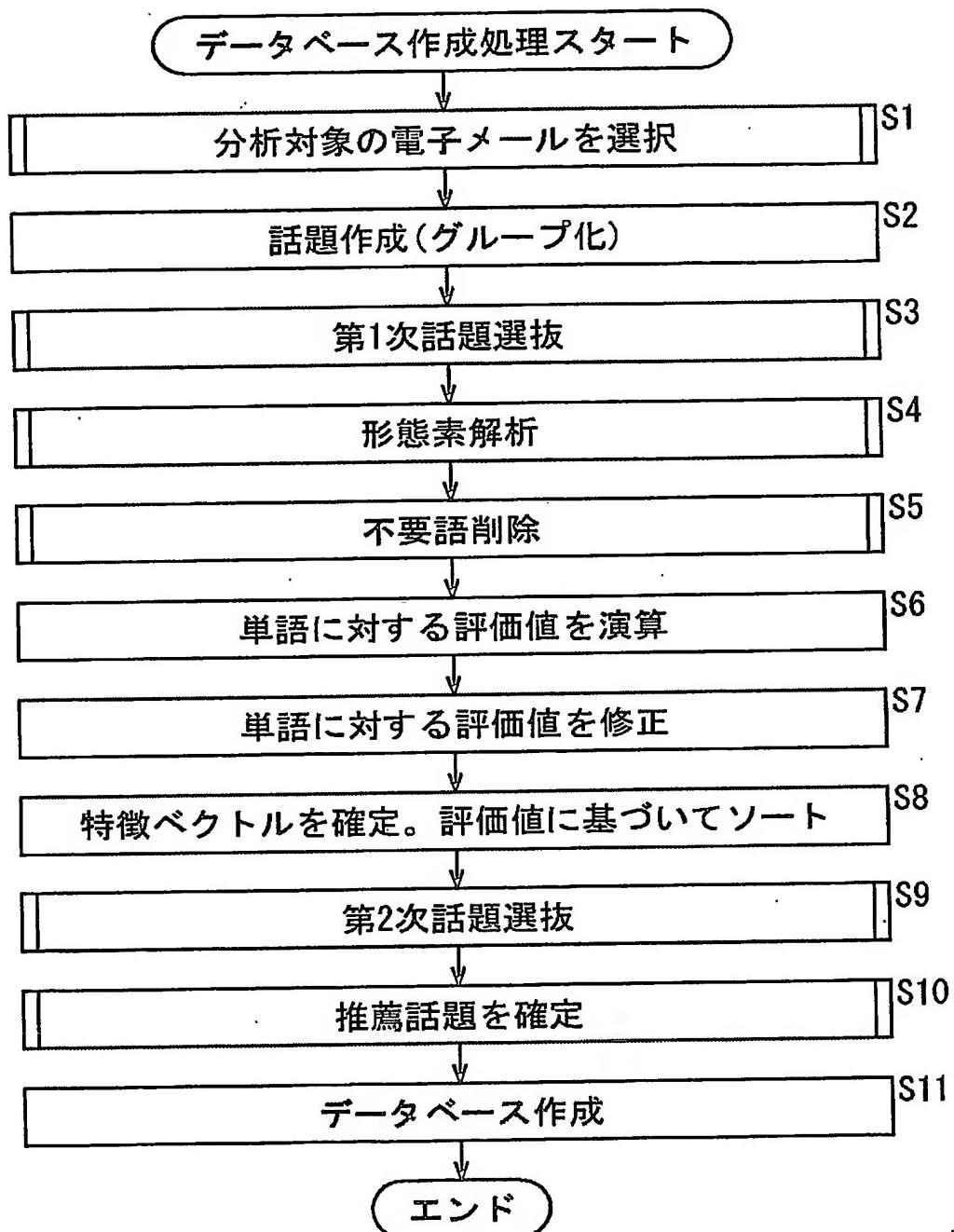
【図6】

図6



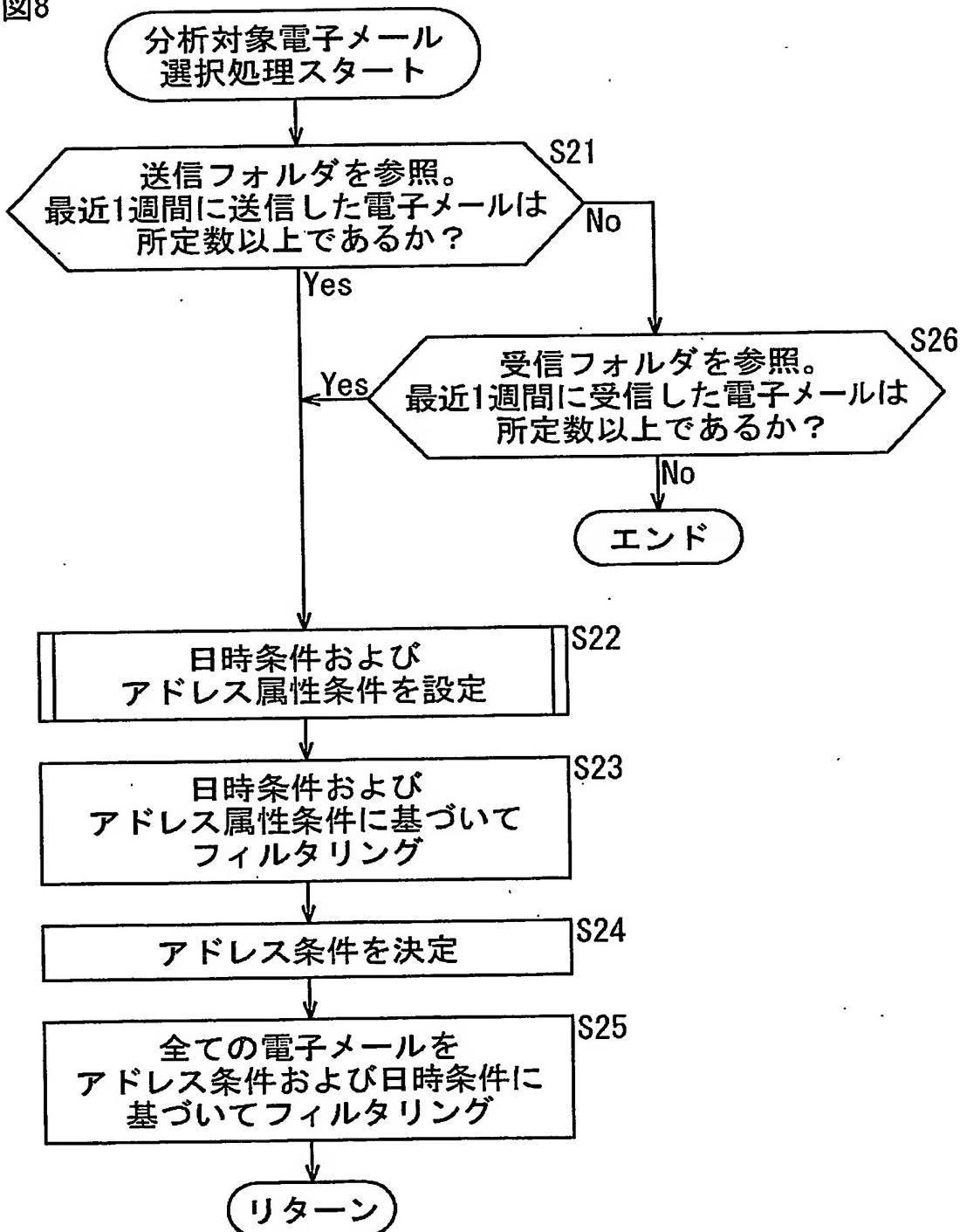
【図7】

図7



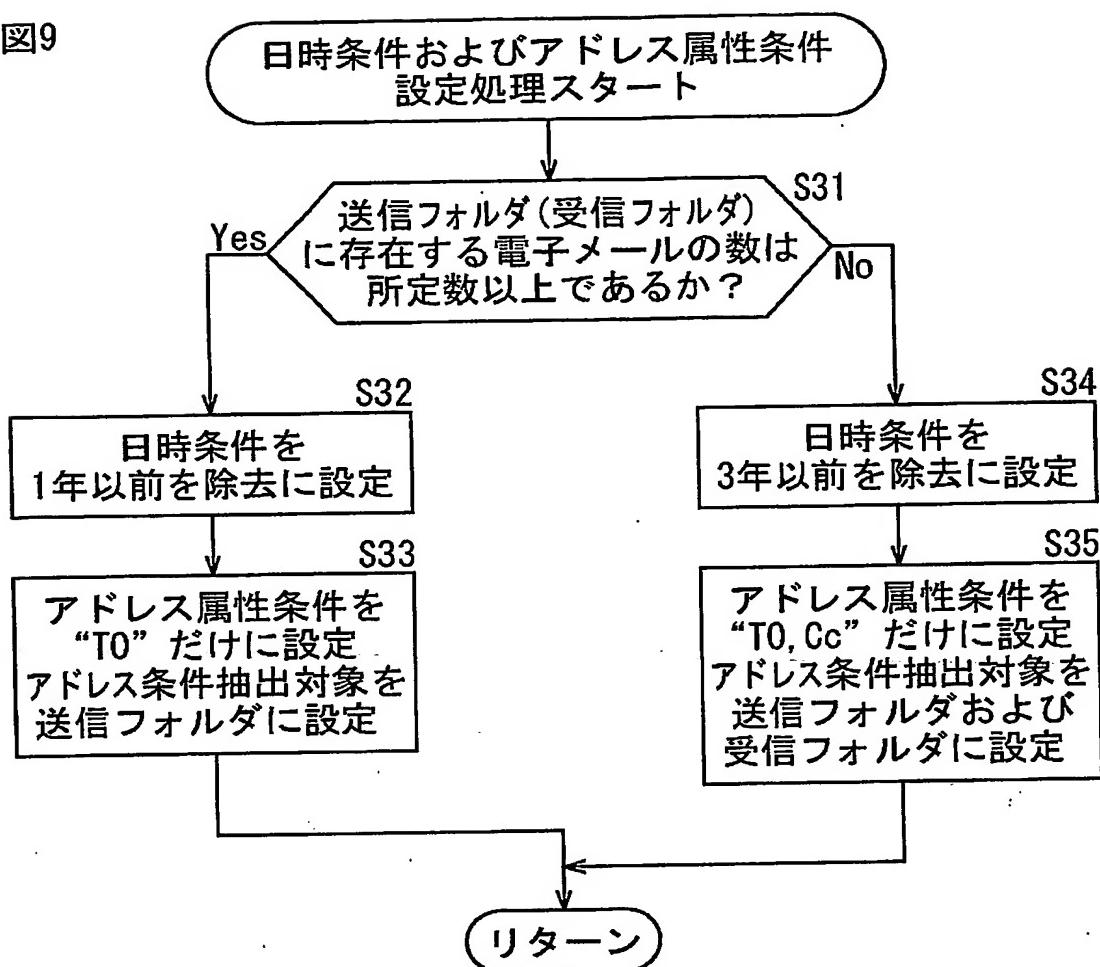
【図8】

図8



【図9】

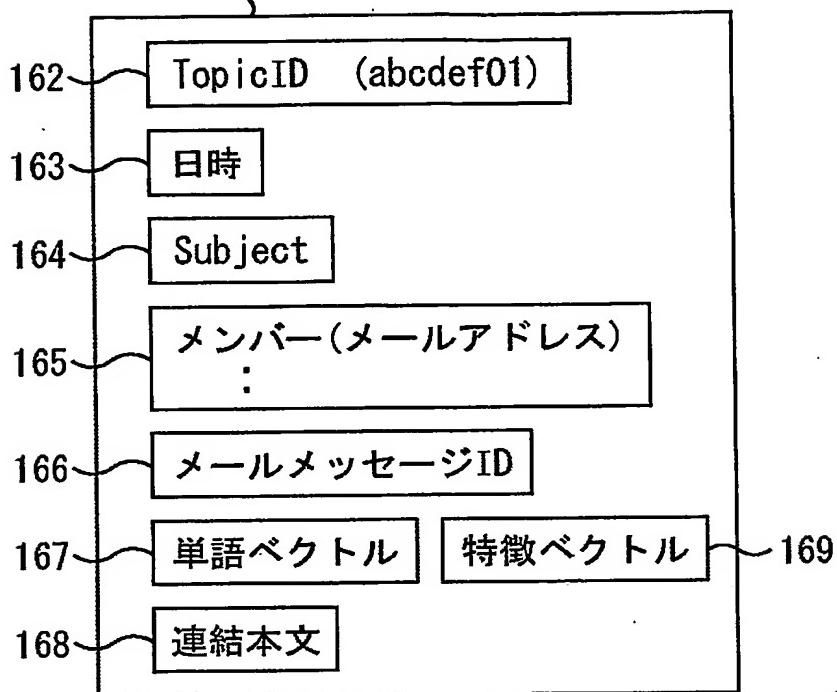
図9



【図10】

図10

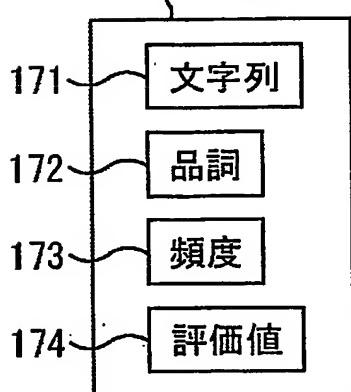
161 話題ファイル



【図11】

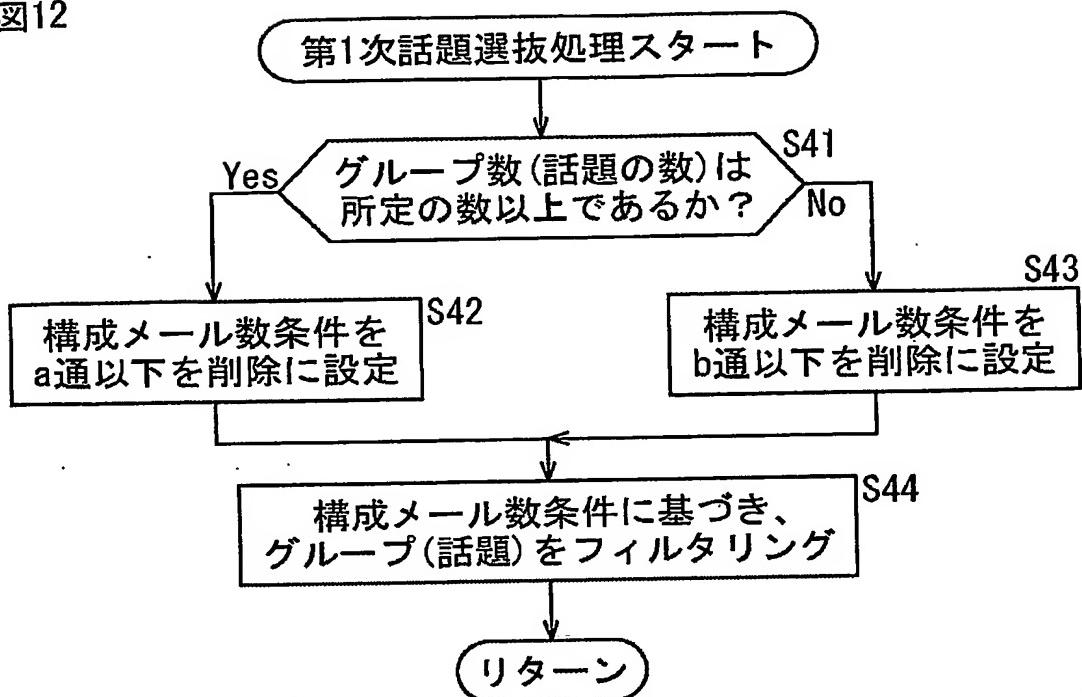
図11

170 単語ベクトル



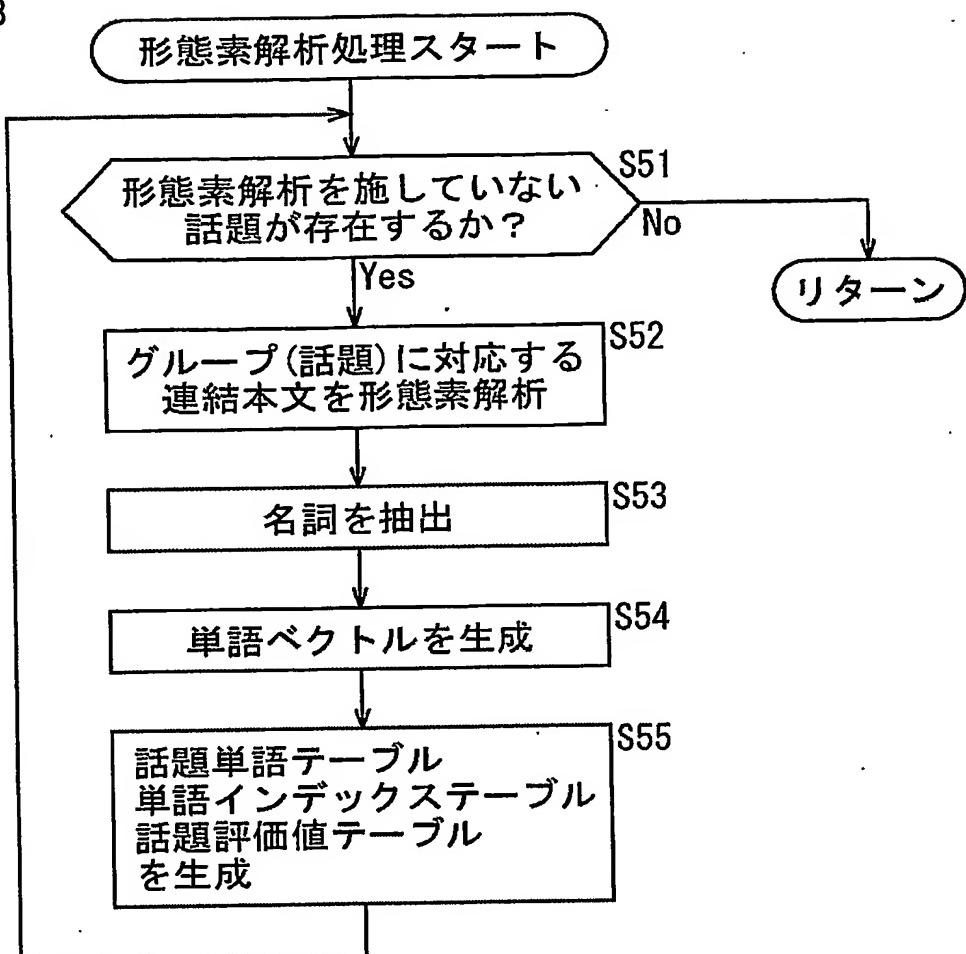
【図12】

図12



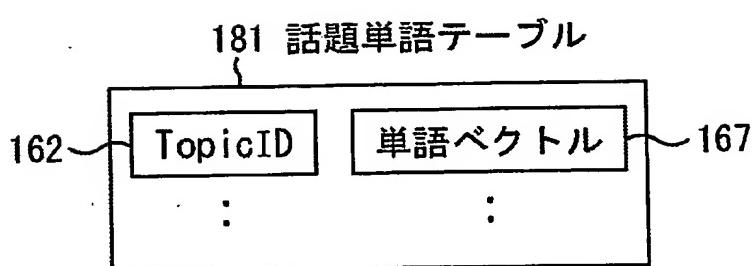
【図13】

図13



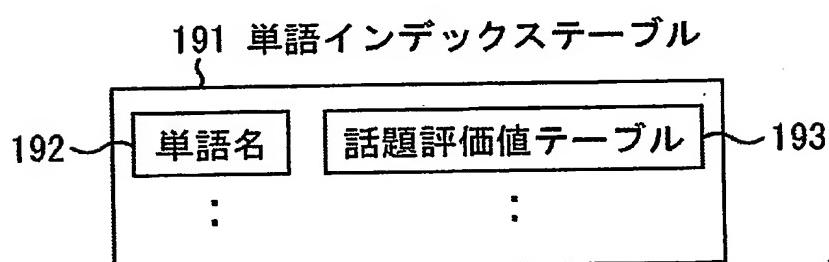
【図14】

図14



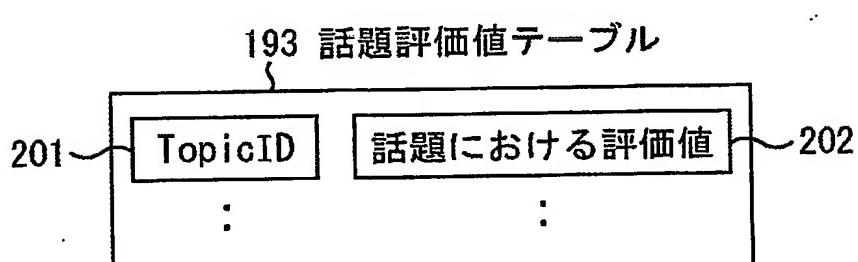
【図15】

図15



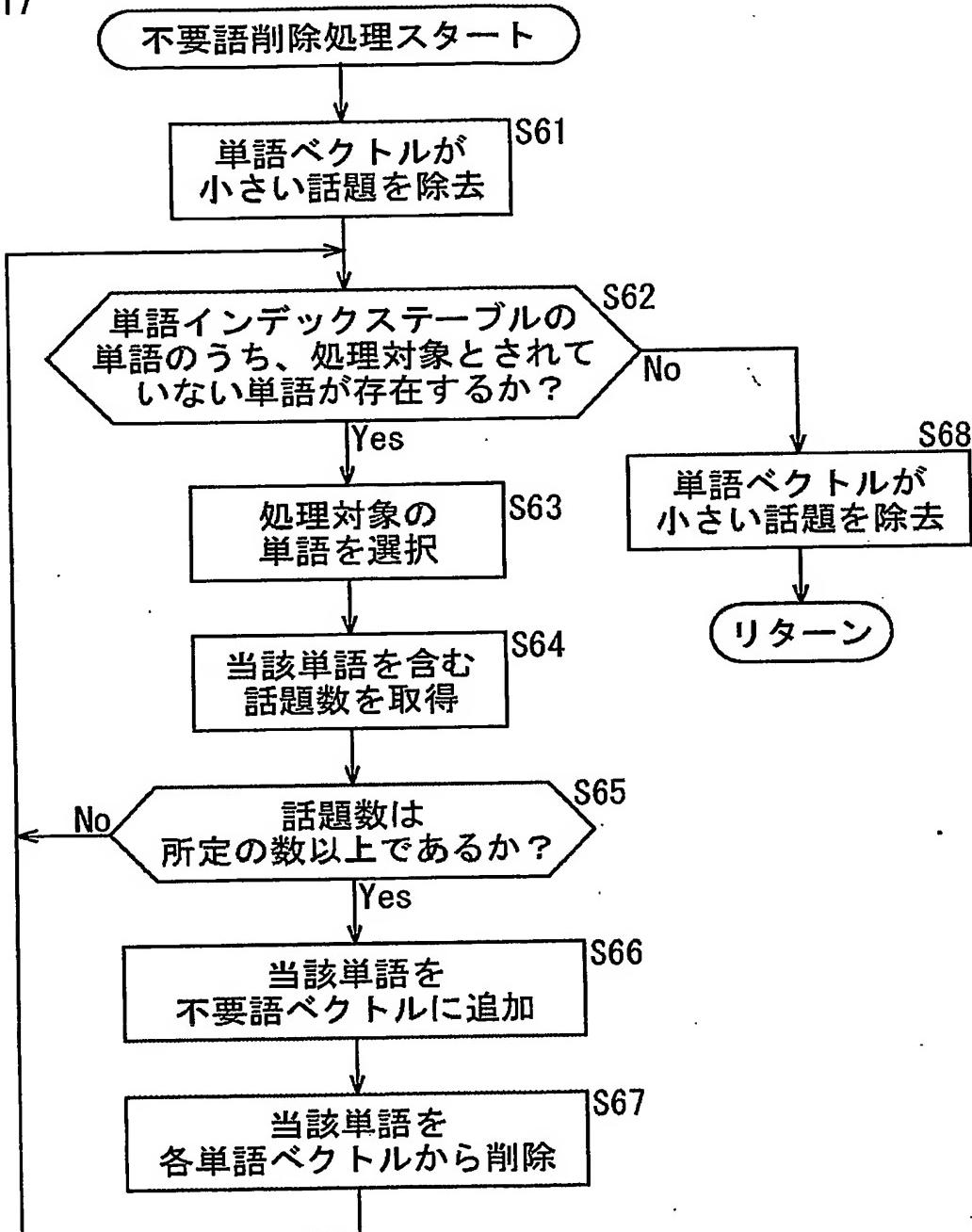
【図16】

図16



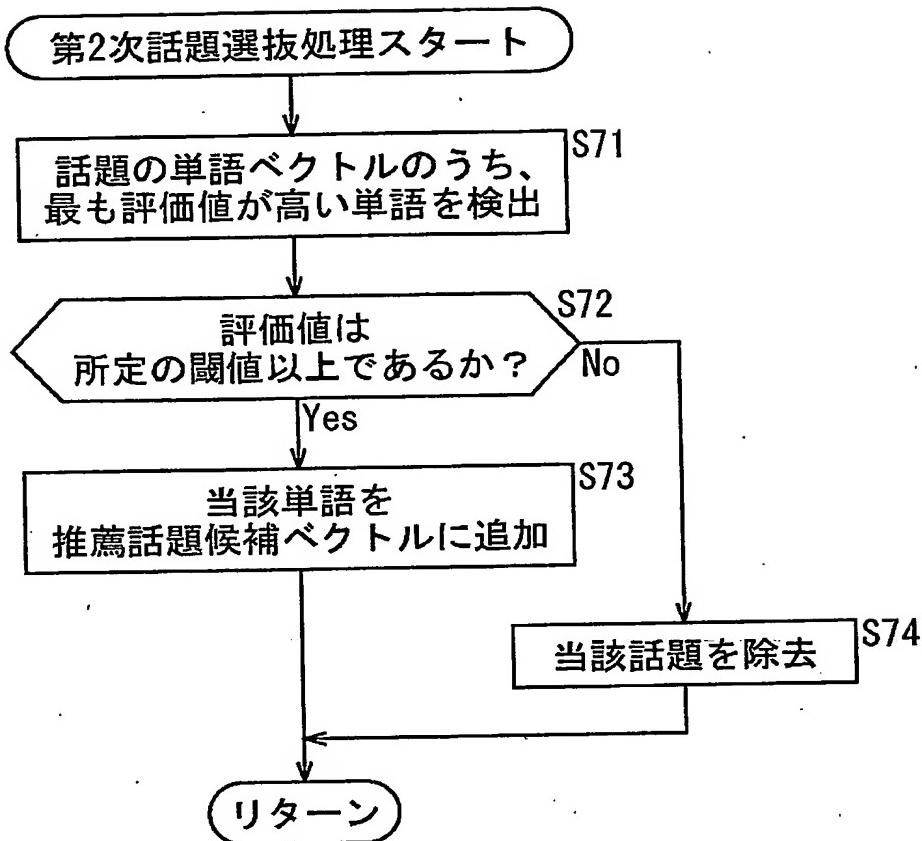
【図17】

図17



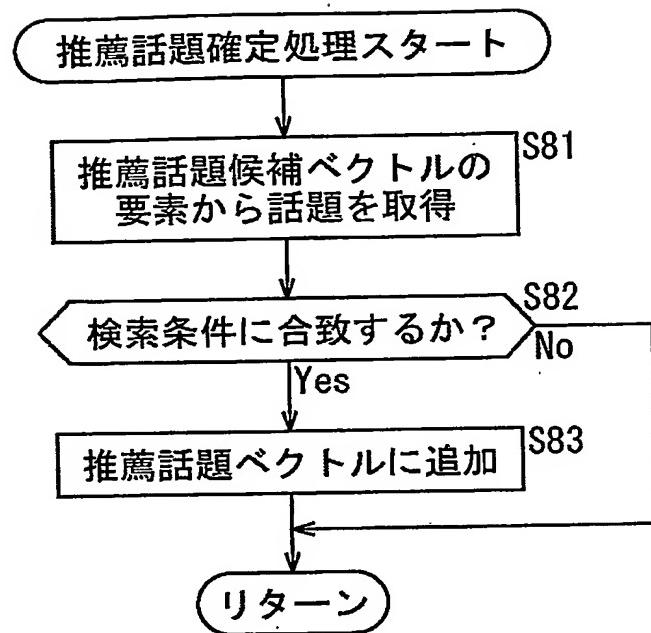
【図18】

図18



【図19】

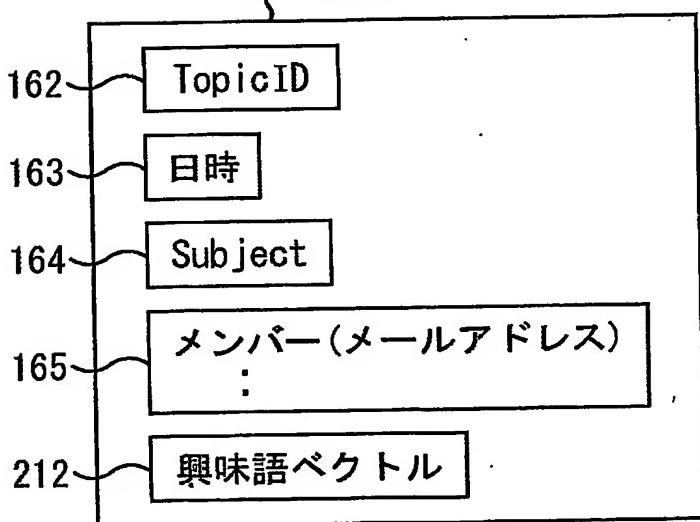
図19



【図20】

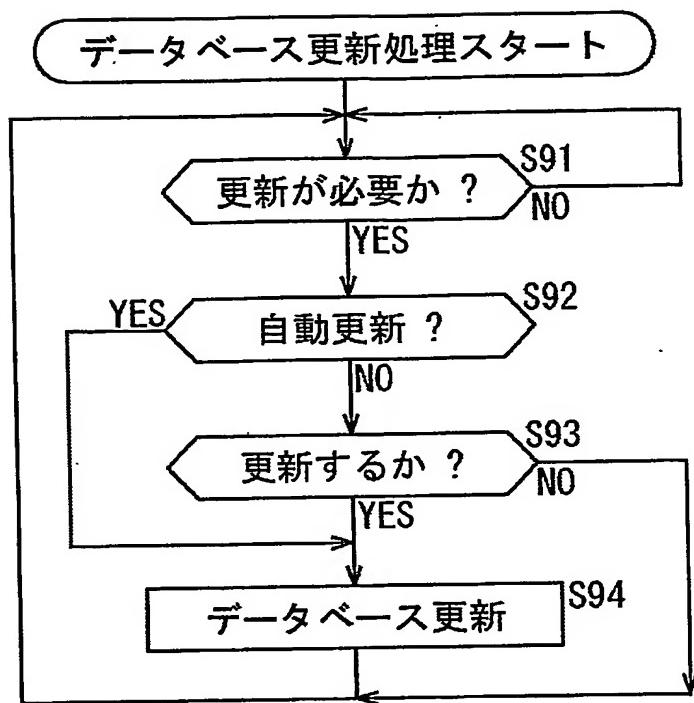
図20

211 興味データ



【図21】

図21



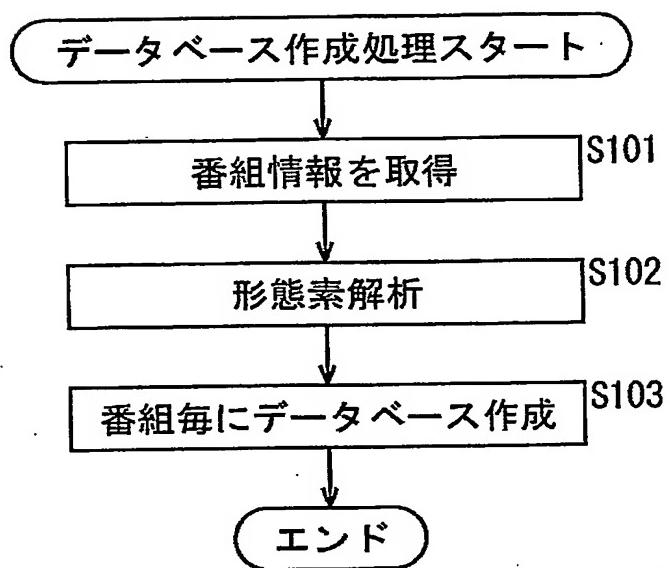
【図22】

図22

- URLが [] 回行われたメールグループは更新必要と判断する
- 推薦URLを使い切ったメールグループは更新必要と判断する
- 最終更新から [] 日経過したメールグループは更新必要と判断する
- 新着メールの分析により、検索語が変更されたメールグループは更新必要と判断する

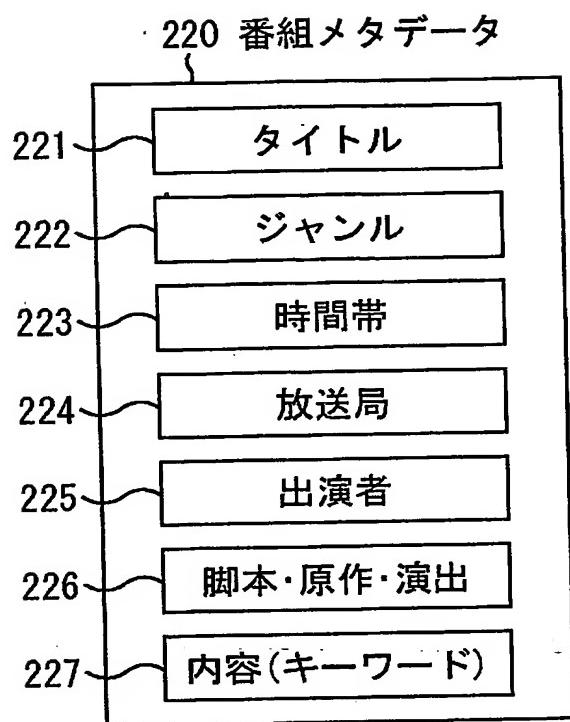
【図23】

図23



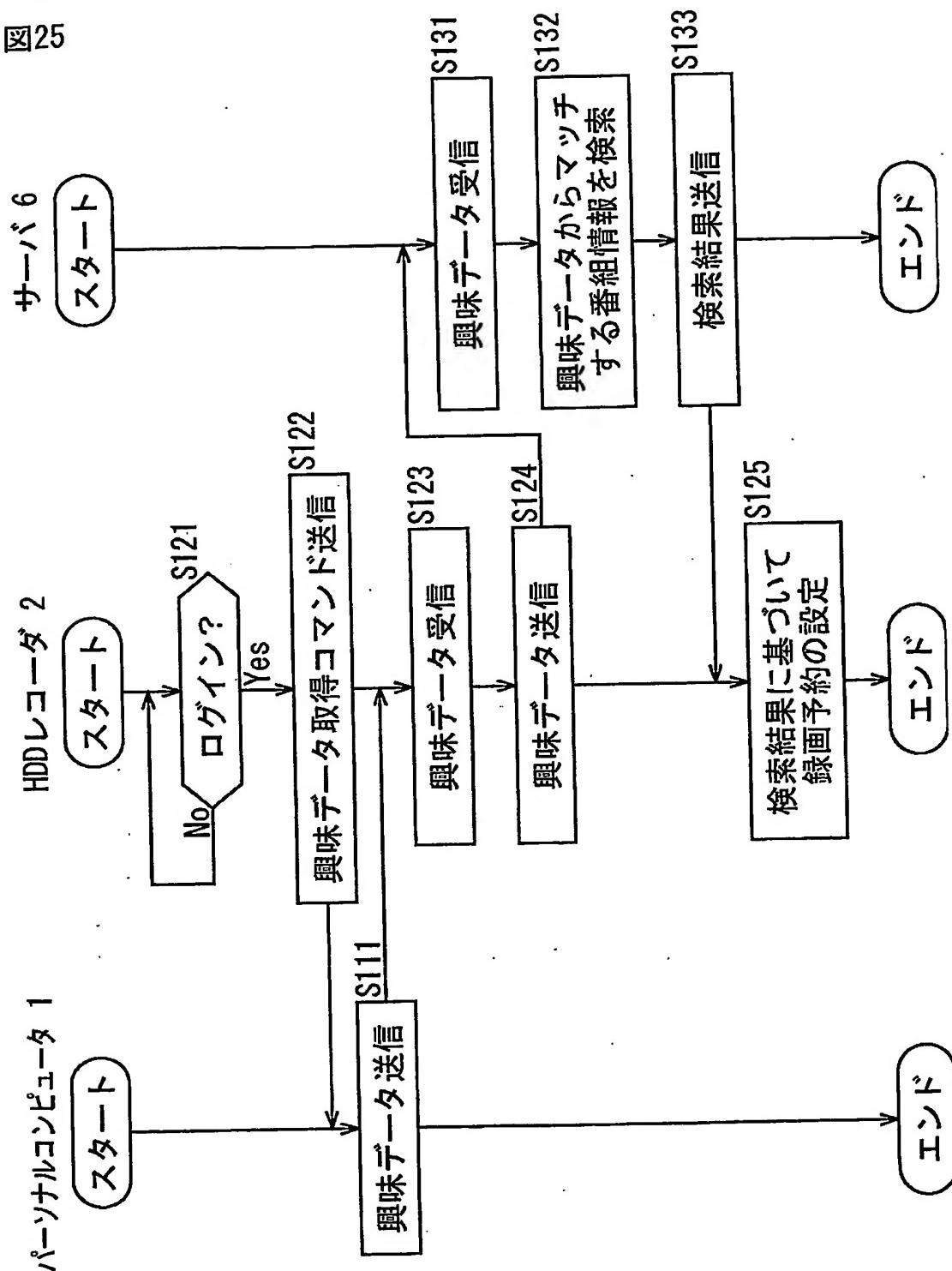
【図24】

図24



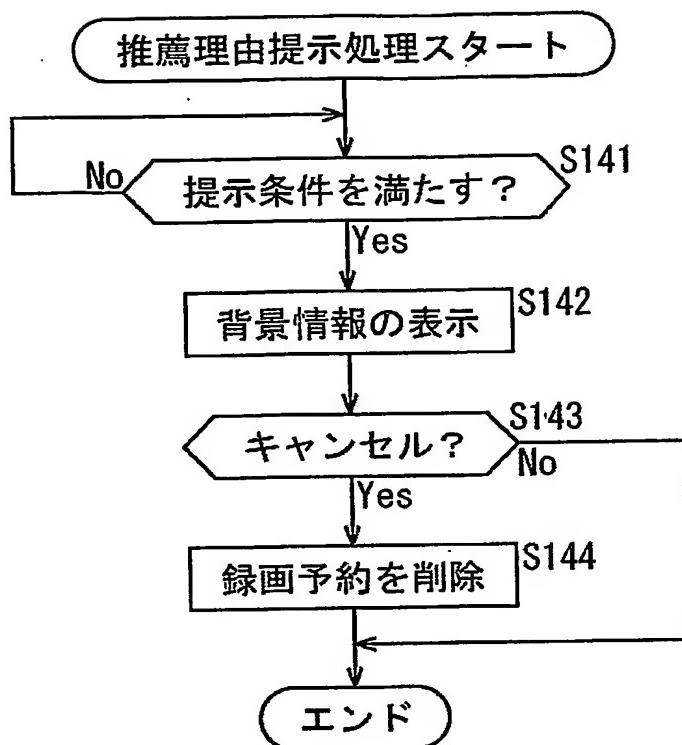
【図25】

図25



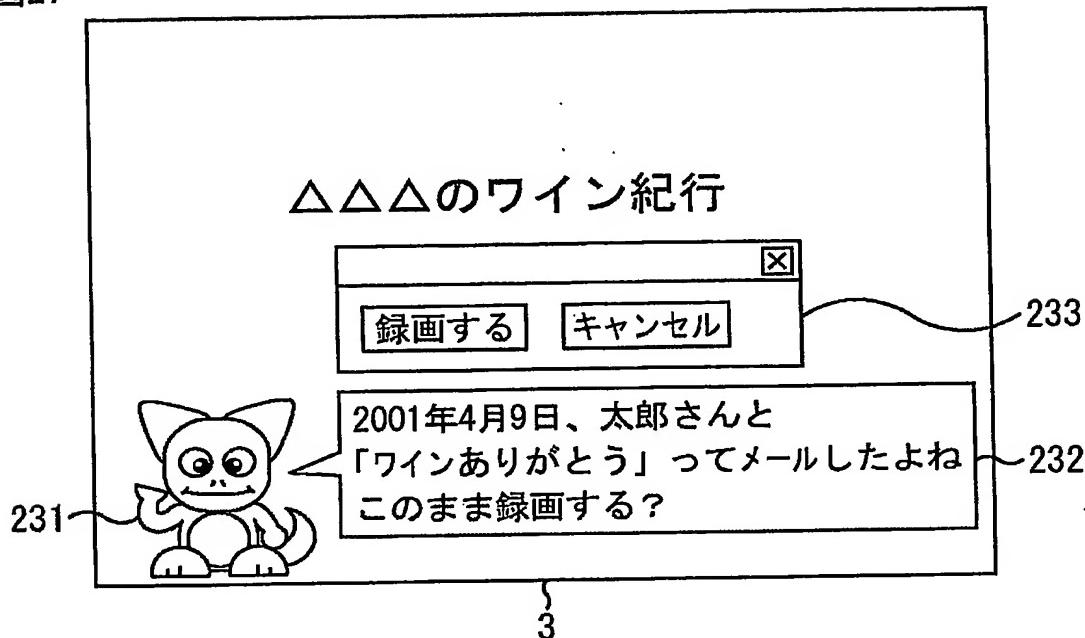
【図26】

図26



【図27】

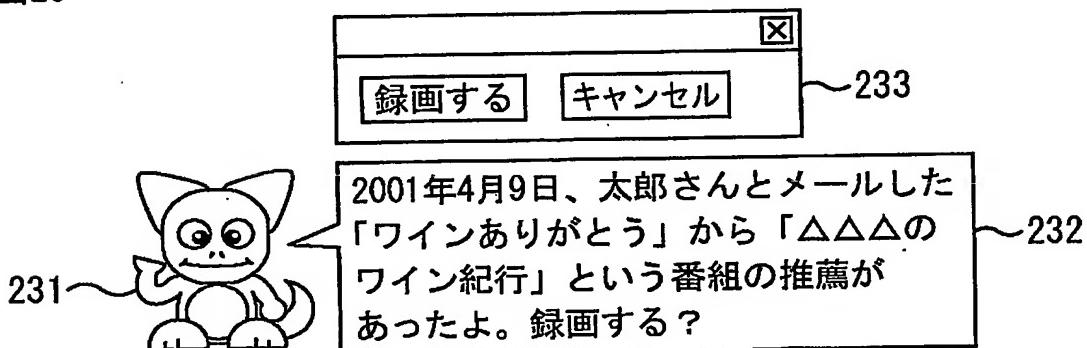
図27



3

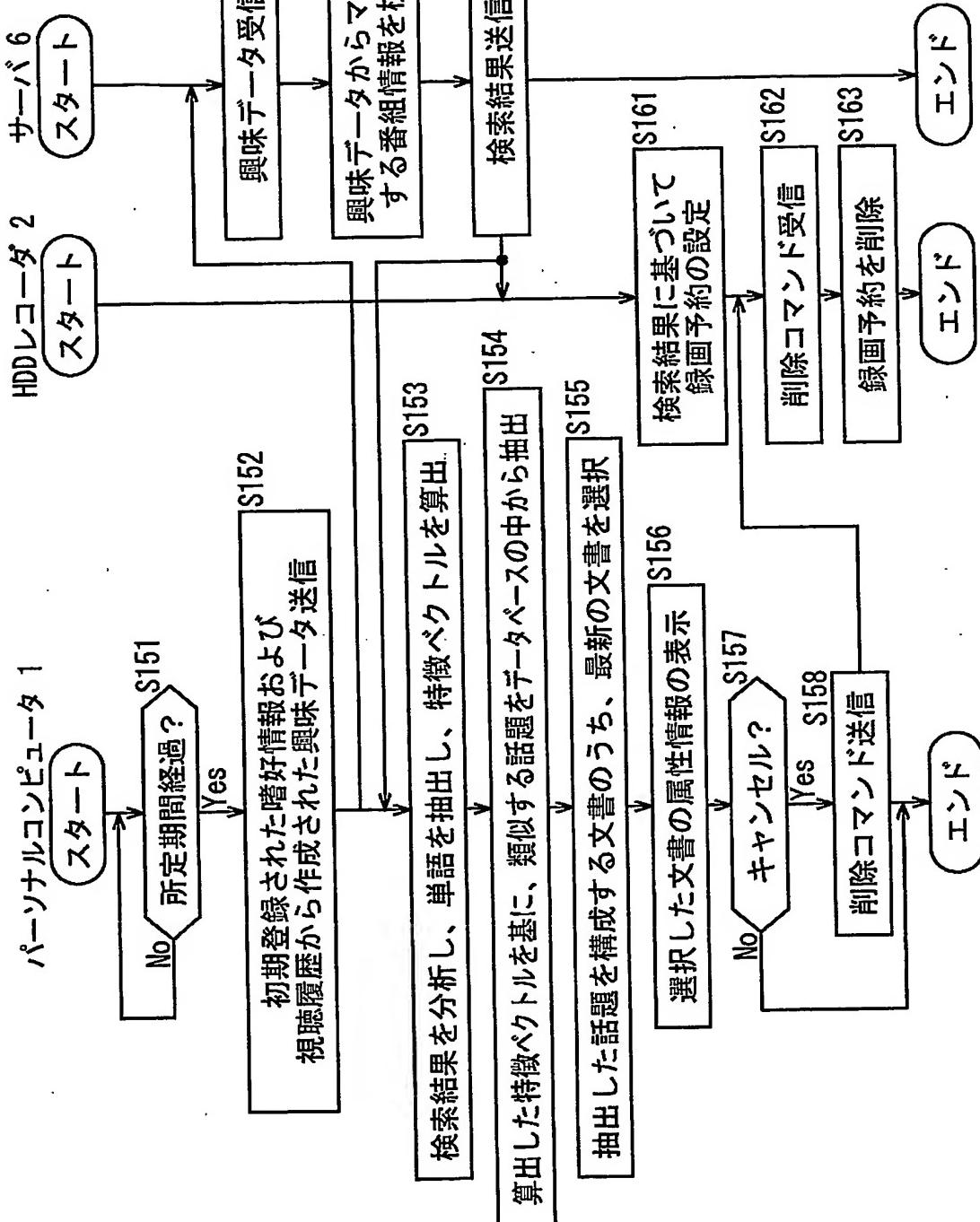
【図28】

図28



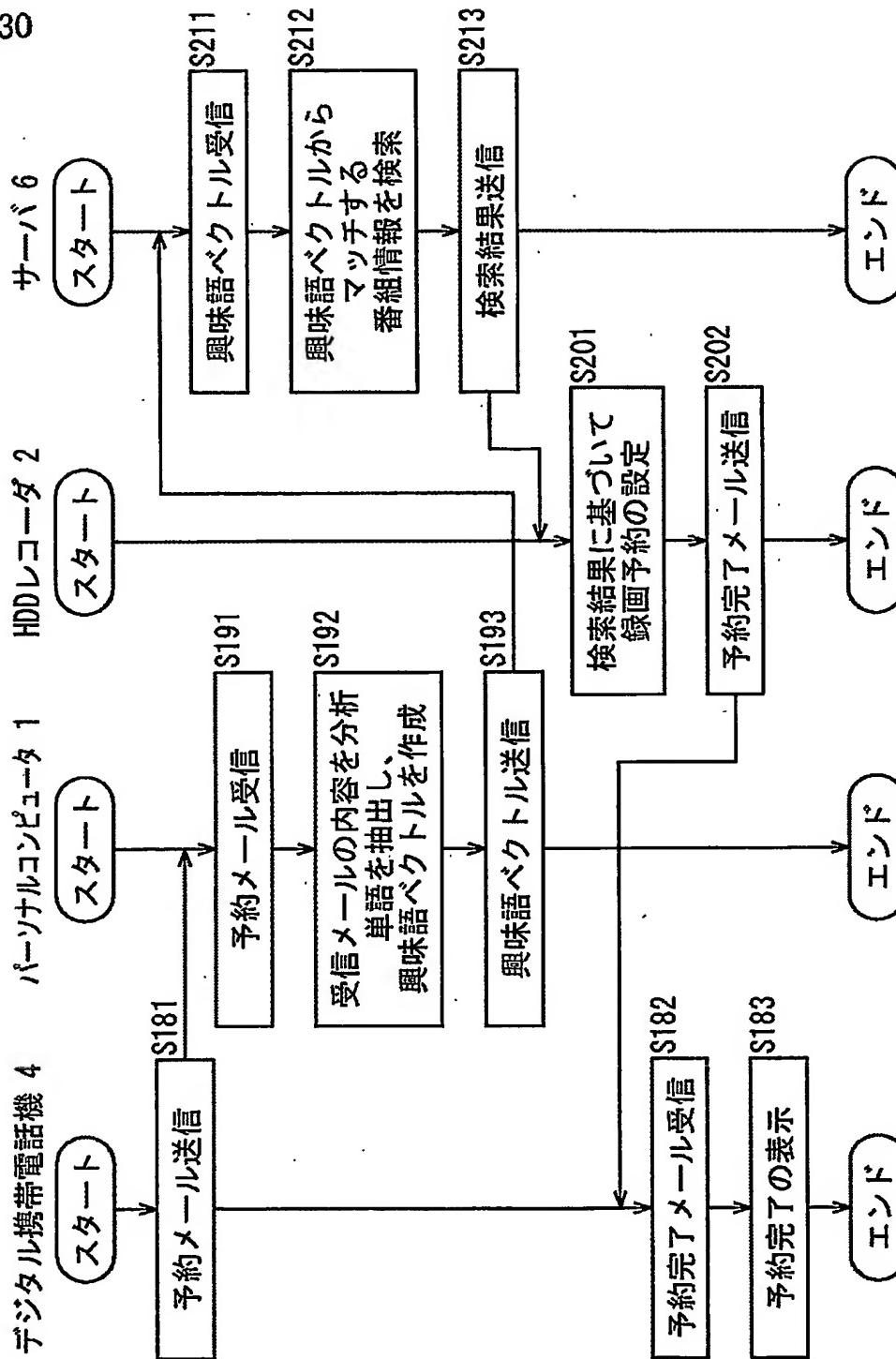
【図29】

図29



【図30】

図30



【図31】

図31

To: * * @aa.co.jp
From: ○○○@bb.co.jp
Subject: 予約メール

今日、音楽番組を録画して。
特にクラシックものは絶対。
あればジャズ番組も。
あとワールドカップ情報も
お願ひ。

241 予約メール

【図32】

図32

To: ○○○@bb.co.jp
From: △△@cc.co.jp
Subject: 予約完了メール

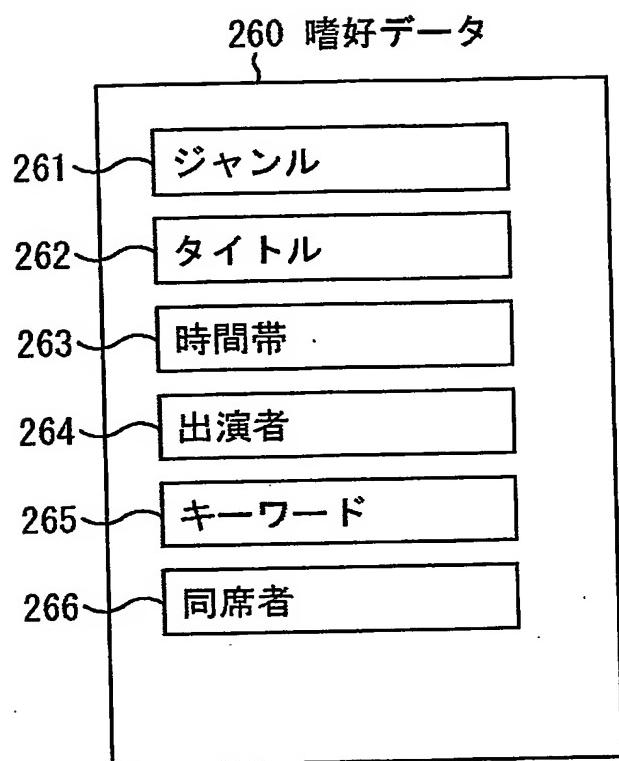
下記の番組を録画しました

- 1 4CH 19:00~20:00
「ワールドカップハイライト」
- 2 3CH 21:00~21:54
「×××クラシック」

251 予約完了メール

【図33】

図33



【書類名】 要約書

【要約】

【課題】 ユーザの興味に対応する番組情報を推薦する。

【解決手段】 パーソナルコンピュータ（以下、PCと記述）は、ユーザが送受信した電子メールの文書からユーザの興味に対応する単語を抽出し、興味データを記録する。S121, S122で、HDDレコーダは、ユーザによりログインされると、興味データの取得を要求する。S111で、PCは、HDDレコーダからの要求に基づいて、ログインされたユーザに対する興味データを送信する。S123, S124で、HDDレコーダは、興味データを受信し、それをサーバに送信する。S131で、サーバは、受信した興味データにマッチする番組情報を検索し、検索結果を送信する。S125で、HDDレコーダは、検索結果に含まれる番組情報に基づいて、番組の録画予約を設定する。本発明は、パーソナルコンピュータにインストールされるプログラムに適用できる。

【選択図】 図25

出願人履歴情報

識別番号 [000002185]

1. 変更年月日 1990年 8月30日

[変更理由] 新規登録

住 所 東京都品川区北品川6丁目7番35号

氏 名 ソニー株式会社